Biochemical Genetics (2019) 57:338–353
https://doi.org/10.1007/s10528-018-9893-2

ORIGINAL ARTICLE

Development of New SSR (Simple Sequence Repeat) Markers for Lentils (*Lens culinaris* Medik.) from Genomic Library Enriched with AG and AC Microsatellites

Melike Bakır[1,2] · **Abdullah Kahraman**[3]

Received: 4 June 2018 / Accepted: 1 November 2018 / Published online: 8 November 2018

Abstract

Simple sequence repeat (SSR) markers are the major molecular tools for genetic and genomic researches that have been extensively developed and used in major crops. However, few are available for lentils (*Lens culinaris* M.), economically an important cool-season legume. The lack of informative simple sequence repeat (SSR) markers in lentil has been a major limitation for lentil molecular breeding studies. Therefore, in order to develop SSR markers for lentil, an enriched genomic libraries for AC and AG repeats were constructed from the *Lens culinaris* cv Kafkas. A total of 350 clones were inquired for the detection of SSRs. Of 350 clones, 68 had SSR motifs. In polymorphism analysis using 53 newly developed SSRs, a total of 144 alleles across 24 lentil cultivars were detected with an average of 4.64 per locus. The average heterozygosity was 0.588 and polymorphism information contents ranged from 0.194 to 0.895 with an average value of 0.520. These newly developed SSRs will constitute useful tools for molecular breeding, mapping, and assessments of genetic diversity and population structure of lentils.

Keywords Lentil · Genomic SSRs · Enriched genomic library · Functional markers · Genetic diversity

✉ Melike Bakır
melikecu@gmail.com

[1] Department of Agricultural Biotechnology, Faculty of Agriculture, Erciyes University, Kayseri 38039, Turkey

[2] Betul-Ziya Eren Genome and Stem Cell Center, Erciyes University, Kayseri 38039, Turkey

[3] Department of Field Crops, Faculty of Agriculture, Harran University, Sanliurfa 63300, Turkey

Introduction

Lentil (*Lens culinaris* Medikus) is one of the oldest domesticated grain legumes. It is an annual, self-pollinated diploid ($2n = 14$) cool-season legume crop with a haploid genome size of 4063 Mbp (Arumuganathan and Earle 1991). Lentil is an important source of dietary protein (22–35%) in both the human nutrition and in animal feed, and it also provides rotational benefits for management of weeds, diseases, and pests, and in many cases offers a profitable, high value crop option for farmers (Hamwieh et al. 2005; Phan et al. 2007). Lentil (*Lens culinaris* Medik. ssp. *culinaris*) is cultivated throughout Europe, Western Asia, the Middle East, North Africa, the Indian subcontinent, North America, and Australia. The archeological records place lentil domestication from its wild progenitor *Lens culinaris* spp. *orientalis* in Syria and Turkey approximately 8500 BC (Cubero 1981). Worldwide lentil production in 2016 was 6.3 million metric tons from an area of 5.48 million ha, the top producers being Canada, India, Turkey, and the United States of America (FAO 2013).

Molecular markers have been used by lentil breeders and geneticists in genetic analysis of lentil (Kumar et al. 2014). Genetic diversity assessment of lentil have been carried out using AFLP, RFLP, RAPD, ISSR, SSR markers (Havey and Muehlbauer 1989; Sharma et al. 1996; Ferguson et al. 1998; Sonnante and Pignone 2001; Toklu et al. 2009; El-Nahas et al. 2011; Alghamdi et al. 2014; Dikshit et al. 2015; Idrissi et al. 2015; Tsanakas et al. 2018). Lentil genetic maps have also made it possible to have a better understanding of lentil genome (Eujayl et al. 1998; Kahraman et al. 2014; Tanyolac et al. 2010; Saha et al. 2013; Ates et al. 2018). However the lack of available molecular markers limits genetic and genomic analysis of lentils as compared to other legumes, limited availability of molecular tools also hinders breeding programs to be carried out to improve lentil cultivars. Therefore, in order to enable breeders to produce varieties with high yield and better quality, efficient molecular tools, like markers, should be developed and used in further breeding programs in lentils.

Among the different types of DNA markers, SSR markers are considered as an important tool for studying genetic diversity, population structure, phylogenetic relationships, construction of frame-work linkage maps, QTL interval mapping, map-based cloning of genes, marker-assisted selection (MAS), etc., thereby aiding in genetic improvement of crop plants (Hendre et al. 2007). SSRs have several genetic advantages, such as high degree of polymorphism, multi-allelic nature, reproducibility, co-dominant inheritance, locus specific, relative abundance, and good genome coverage (Powell et al. 1996).

The first genomic library formed to develop SSR markers in lentils was performed in ILL5588 cultivar by using *SauIII* restriction enzyme (Hamwieh et al. 2009). It was reported that 371 (0.18%) of 200,000 clones screened through GT, GA, GC, GAA, TA, TAA repeats contained microsatellites and 243 (65.4%) of them were sequenceable. Of these 243 sequenced clones, 173 (71.2%) contained SSR motifs. Verma et al. (2014) used Precoz lentil cultivar and sequenced 514 clones from genomic libraries enriched with GA/CT repeats and reported that

375 (72.9%) them contained three or more SSR motifs. Andeden et al. (2015) worked with Karacadag lentil genotype and inquired 432 clones from genomic libraries enriched with CA, GA, AAC, and ATG repeats and encountered SSR motifs in 360 (83.3%) clones.

So far, the number of available polymorphic genomic SSR markers for lentils is only 244 (Hamwieh et al. 2005, 2009; Verma et al. 2014; Andeden et al. 2015). Therefore, the objective of the present study was to develop a new set of SSR markers from the microsatellite-enriched genomic library of lentils and to determine polymorphism rate of these markers for the analysis of genetic diversity in Turkish lentil genotypes. This new genomic resource of SSR markers would provide significant contributions in molecular breeding of lentils.

Material and Methods

Plant Material and DNA Extraction

A total of 23 lentil cultivars (Firat-87, Tigris, Seyran-96, Cagil, Altıntoprak, Yerli Kırmızı, Kafkas, Ankara Yesili, Bozok, Ceren, Gumrah, Karagul, Yusufhan, Ali Dayı, Ciftci, Meyveci-2001, Emre-20, Sultani-1, Sazak-91, Kayi-91, Ozbek, Malazgirt-89, Erzurum-89) developed in different research centers of Turkey and black lentil cultivar Beluga were used for amplification of developed SSR markers and to determine polymorphism ratios. Genomic DNA was isolated from fresh, young leaves of all accessions according to the protocol described by Lefort et al. (1998) with minor modifications. The quality and quantity of the extracted DNA were determined in NanoDrop® ND-1000 Spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA) and agarose gel electrophoresis (0.8%).

Construction of a Genomic Library Enriched for the AG and AC Microsatellite

A genomic library of *L. culinaris* cv. Kafkas, enriched for the AG and AC motif was constructed using a modified protocol of Techen et al. (2010). For this purpose, the biotinylated $(AG)_{12}$ and $(AC)_{12}$ oligoprobe and magnetic beads coated with streptavidin were used following the hybridization-based capture technique. Briefly, nuclear DNA of lentil (cv. Kafkas) was restricted with a combination of RsaI+AluI+HaeIII (NEB) in the same reaction. Genomic DNA fragments were A-tailed and ligated to specific adaptors [blunt end primers SSRLIBF3 (5′-CGGGAGAGCAAGGAA GGAGT-3′) and SSRLIBR3 (5′-Phos CTCCTTCCTTGCTCTCCCGAAAA-3′)]. Adapter-specific primers SSRLIBF3 were used to amplify the adaptor-ligated DNA fragments. The amplified products were hybridized with the biotinylated microsatellite oligo AG and AC in the same reactions at 50 °C (depending on the Tm of the oligo) for 4 h. Streptavidin-coated magnetic beads (Invitrogen, Dynabead M-280) were used to capture DNA fragments hybridized with the AG and AC-rich biotinylated fragments according to the manufacturer's instructions. After binding, the beads were washed first with 2XSSC, then with 0.5XSSC both at room temperature

and finally with 0.5XSSC at 50 °C for 5 min. Elution of the single strand DNA from the biotinylated oligos was done twice with 60 µl of MQ water at 96 °C for 10 min. and amplified with the SSRLIBF3 primer. The PCR products were cloned into T–A vector TOPO4 (Invitrogen, Carlsbad, CA, USA) and transformed into TOP10 cells (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions. The recombinant colonies were picked and prepared in glycerol stocks. Colony PCR was done to identify colonies containing microsatellites with M13 Forward and Reverse primers and microsatellite oligo AG and AC primers (Bloor et al. 2001). The PCR products were analyzed in 2% agarose gel electrophoresis. The PCR products with two or more bands indicated that the plasmid contains a microsatellite-containing insert.

Analysis of Microsatellite-Containing Sequences and Primer design

Positive plasmids were amplified using the TempliPhi Amplification kit (GE Healthcare, Wauwatosa, WI, USA) and sequenced using BigDye Terminator v3.1 Cycle Sequencing Kit in the Applied Biosystems Prism 3500 Genetic Analyzer System (Applied Biosystems, Foster City, CA, USA). The vector sequence was removed by using the Vecscreen (https://www.ncbi.nlm.nih.gov/tools/vecscreen/) program. To remove redundancy, CAP3 program (https://doua.prabi.fr/software/cap3) (Huang and Madan 1999) was used and the location of microsatellite repeats was determined with SSRIT (https://archive.gramene.org/db/markers/ssrtool) program (Temnykh et al. 2001). Duplicated sequences were identified with BioEdit (Hall 1999) program and they were removed. At least 5 primers were designed for dinucleotide repeats, at least 3 primers were designed for trinucleotide repeats and more primers were designed for the sequences containing the greater number of repeats by using Primer3 (https://bioinfo.ut.ee/primer3-0.4.0/) (Koressaar and Remm 2007; Untergasser et al. 2012) and BatchPrimer3 (https://probes.pw.usda.gov/batchprimer3/) (You et al. 2008) programs.

PCR Amplification of Microsatellites and Genetic Diversity Analysis

In order to validate the markers developed in this study, PCR amplifications (15 µl) were performed with 90 ng genomic DNA, 10 µM of each primer, 0.2 mM of each dNTPs, 1X DreamTaq Green Buffer (includes $MgCl_2$ at a concentration of 2 mM) (Thermo Scientific, Waltham, MA, USA), and 0.5U DreamTaq DNA Polymerase (Thermo Scientific, Waltham, MA, USA). All PCR reactions were performed in the Bio-Rad thermocycler. The amplification program consisted of an initial step of 3 min at 94 °C, followed by 35 cycles of 1 min at 94 °C, 1 min at 50–66 °C, 2 min at 72 °C, and a final extension at 72 °C for 10 min. The amplified products were analyzed through 3% agarose gel electrophoresis.

Genetic diversity analysis was performed by using M13-tailed primer according to the methods described by Schuelke (2000) in 24 lentil genotypes. A tail (M13 universal sequence (−21), TGTAAAACGACGGCCAGT) was added to the 5′ end of each forward primers. PCR amplifications were performed in 15 µl reaction mixture

containing 90 ng genomic DNA, 0.1 µM of each SSR primer, 0.1 µM labeled M13 (−21) universal primer, 0.2 mM of each dNTPs, 1X DreamTaq Green Buffer (includes $MgCl_2$ at a concentration of 2 mM) (Thermo Scientific, Waltham, MA, USA) and 0.5U DreamTaq DNA Polymerase (Thermo Scientific, Waltham, MA, USA). The amplification program consisted of an initial step of 3 min at 94 °C, followed by 35 cycles of 1 min at 94 °C, 1 min at 50–66 °C, 2 min at 72 °C, followed by 8 cycles of 1 min at 94 °C, 1 min at 53°C, 2 min at 72 °C, and a final extension at 72 °C for 10 min. The M13 (−21) primer was 5′-fluorescently tagged with HEX, 6-FAM or ROX to facilitate multiplexing. A set of three PCR products (0.5 µl each) was mixed with 0.5 µl GeneScan-600 LIZ size standards (Applied Biosystems, Foster City, CA, USA) and 9.5 µl Hi-Di™ formamide (Applied Biosystems, Foster City, CA, USA) and denatured at 95 °C for 5 min, chilled on ice and electrophoresed on the Applied Biosystems Prism 3500 Genetic Analyzer System (Applied Biosystems, Foster City, CA, USA). GENEMAPPER software v5.0 (Applied Biosystems, Foster City, CA, USA) was used to determine fragment size.

Data Analysis

Microsatellite diversity analyses were carried out at the locus level in a given species. For each locus, the expected heterozygosity (H_e), observed heterozygosity (H_o) and polymorphism information content (PIC) (Nei 1973) were calculated with PowerMarker V3.025 software (Liu and Muse 2005). The UPGMA (unweighted pair-group method using arithmetic average) were used to construct and draw a dendrogram from the genetic similarity matrix by using the MEGA6 (Tamura et al. 2007) and PowerMarker software programs. Bootstrap analyses with 100 replicates were performed and a consensus tree was obtained to measure the confidence levels for the clusters.

Results

Isolation and Characterization of Microsatellites

A total of 350 clones were inquired for Kafkas lentil cultivar from the libraries enriched by using the AG and AC repeat motifs within the same reaction. These clones were screened with colony PCR reaction containing AG and AC repeat motifs and 68 of them contained repeat regions. When these 68 clones with repeat sections were analyzed through sequencing, it was observed that 53 of them were identified as appropriate for primer design. In clones without a primer design, number of repeat motifs was identified as 4 and less for dinucleotide repeats. The 53 sequences with primer design contained a total of 134 SSR motifs (Table 1). Among the identified SSR motifs, GA/CT motif was the most frequent one (62.6%). The other motifs were identified as AG/TC, GT/CA, AC/TG, CTT/GAA, AGA/TCT (respectively with 23.8%, 6.8%, 4.5%, 1.6%, 0.7%). Microsatellite repeats mostly contained dinucleotide repeats, slightly contained trinucleotide repeats and generally

Table 1 Newly developed lentil SSR primers and information about these primers

Primers	5′→3′	Microsatellite motifs	Expected size (bp)	Temperature (°C)
Lc_MCu1_F	AGAACCGATCGTGTTTGAGC	$(GA)_4...(GA)_3...(GA)_3...(GA)_4...(GA)_{6imperf}$	250	58
Lc_MCu1_R	GGGTGGAAAACCCTAATTCC			
Lc_MCu2_F	CTTCTGCAGCCAGCAACTC	$(CTT)_3...(CT)_4TTT(CT)_3...(CT)_{7imperf}$	241	60
Lc_MCu2_R	CAGCAGCATAAAAACGAGGA			
Lc_MCu3_F	AGGAGTACGGAACCGGAGAT	$(GA)_3T(GA)_3...(GA)_4...(GA)_3...(GA)_3...(GA)_{4imperf}$	250	60
Lc_MCu3_R	CCAGTCGTTTCATTCTCTTCC			
Lc_MCu4_F	CCGTCGGTCGTTTCATTCTA	$(CT)_6...(CT)_4...(TC)_{5imperf}$	244	58
Lc_MCu4_R	AGAGAATGACGCGGTTGG			
Lc_MCu5a_F	TTTCAGCAAAAATAGGGTTTCA	$(GA)_7...(GA)_{7imperf}$	151	60
Lc_MCu5a_R	GGTTGAAAATTGGAAAGGGATT			
Lc_MCu6_F	TTACTGTCGCAACCGTTAAG	$(GA)_4...(GA)_{5imperf}$	170	56
Lc_MCu6_R	CCCTCATTCTCCCTCATTCA			
Lc_MCu7_F	AGAACCGATCGTGTTTGAGC	$(GA)_4...(GA)_4...(GA)_4...(GA)_{4imperf}$	247	58
Lc_MCu7_R	GGGTGGAAAACCCTAATTCC			
Lc_MCu8_F	AGAAGAAATCGCAACCCAAA	$(GAA)_5...(GA)_4...(GA)_4...(GA)_{4imperf}$	192	58
Lc_MCu8_R	GGCGGCAATACTCTCTTCTC			
Lc_MCu9_F	TCCGAGTGGTGTGTGATGTT	$(TG)_{5perf}$	151	58
Lc_MCu9_R	GACAAGGCGCTATTGGTGAT			
Lc_MCu10_F	GCACTCCAATGCTCAAGTCA	$(GT)_{12perf}$	250	60
Lc_MCu10_R	GAGTCCAACCACCTGGAAGT			
Lc_MCu11_F	GAGTGGGAAGGAGACCACAA	$(CT)_5CG(CT)_{3imperf}$	160	54
Lc_MCu11_R	CGTGGTCAGGAGAGGGAATA			
Lc_MCu12b_F	TTGAGTGAATGGTGAGGGAAGAGA	$(GA)_{9perf}$	171	58
Lc_MCu12b_R	GCAACGAAGGAGTGGGGAGA			
Lc_MCu13a_F	AAAGTATAGGAACCCACCAAC	$(TA)_4...(AC)_{7imperf}$	154	58

Table 1 (continued)

Primers	5′→3′	Microsatellite motifs	Expected size (bp)	Temperature (°C)
Lc_MCu13a_R	AAGTGAGAGTTCACACGACTG			
Lc_MCu14b_F	CCGGTTTGGTCAACGGAAAG	$(GA)_{12perf}$	126	62
Lc_MCu14b_R	ACAGTGTGACGGCGTGAAGG			
Lc_MCu15_F	TTTTTGGTTTAATGTTGGAACG	$(AG)_{10perf}$	163	58
Lc_MCu15_R	AGGAGTCCTCCTTCTCCTTTG			
Lc_MCu16b_F	CGCGAGAGAGAAAGCCCATC	$(GA)_4...(GA)_5...(GA)_7...(GA)_{9imperf}$	115	64
Lc_MCu16b_R	GCGGCGGCAATACTCTCTTC			
Lc_MCu17_F	GAAAGACAAAGAACGTGATAGAAGG	$(GA)_7...(AG)_4...(AG)_{4imperf}$	157	58
Lc_MCu17_R	TGACCGTTGTTCCCAAATTC			
Lc_MCu18_F	TGGGAATTCTTTCGTTGGAG	$(GA)_4TT(AG)_{11imperf}$	235	62
Lc_MCu18_R	TACCTTGGAGGCATCCTCAT			
Lc_MCu19_F	TGTGAACAATTTAGCAAGCAA	$(AG)_{11perf}$	240	56
Lc_MCu19_R	AGGCAAGAACGATTGGAGAA			
Lc_MCu20_F	TTGATGGCTTGTTCCATTCA	$(GT)_{5perf}$	237	56
Lc_MCu20_R	TAGGGGCTCAGCTCATCATC			
Lc_MCu21_F	AGAACCGATCGTGTTTGAGC	$(GA)_5...(GA)_4...(GA)_{6imperf}$	247	58
Lc_MCu21_R	GGGTGGAAAACCCTAATTCC			
Lc_MCu22_F	TCATGTTGGGACTTTGGTCA	$(GT)_5G(GT)_{3imperf}$	239	58
Lc_MCu22_R	GTCATGGAGAGGACCCAAAA			
Lc_MCu23a_F	AGTCAAAGAAGAGGGAAATTG	$(AG)_4...(GA)_5...(AG)_{4imperf}$	128	56
Lc_MCu23a_R	GAGAGCAACGAAGGAGTTC			
Lc_MCu24_F	TCCCAATTGACCACTCCATA	$(CG)_6...(TC)_{10imperf}$	245	58
Lc_MCu24_R	CAAACGCAAGCATCAACATC			

Table 1 (continued)

Primers	5′→3′	Microsatellite motifs	Expected size (bp)	Temperature (°C)
Lc_MCu25_F	ATGCGGTAAGGTGGTGTCTC	$(TC)_5...(CT)_3T(CT)_{2imperf}$	153	60
Lc_MCu25_R	GACCAAGATGATGCAGAACG			
Lc_MCu26a_F	TGGGGGATGATGATGCAGAG	$(GA)_4...(GA)_{6imperf}$	128	58
Lc_MCu26a_R	AAAACTCCCAGGCGCAAAAA			
Lc_MCu27a_F	AGAGGCCGTGAGCAGTGAGG	$(GA)_{5perf}$	301	60
Lc_MCu27a_R	TCCTTCCACTCCAAACCCACA			
Lc_MCu28_F	TCGATTTGGGGGTCTATGTT	$(GT)_3TGG(GT)_{4imperf}$	167	56
Lc_MCu28_R	ATTTCATCCCCAAACACCAA			
Lc_MCu29a_F	TGTCGATTTCAGAGATGAGTTA	$(TC)_{7perf}$	300	55
Lc_MCu29a_R	AGAGCAACGAAGGAGTAGG			
Lc_MCu30_F	ATGAGGTGCGTTGAGATGGT	$(AG)_{11perf}$	159	60
Lc_MCu30_R	GTCCCAACCCCTATCCAACT			
Lc_MCu31_F	GGGTGGAAAACCCTAATTCC	$(CT)_5...(CT)_4...(TC)_{5imperf}$	245	55
Lc_MCu31_R	AGAACCGATCGTGTTTGAGC			
Lc_MCu32a_F	TCCAACAGTTAGGAGTGTACTG	$(TC)_4T(TC)_3...(TC)_{5imperf}$	193	58
Lc_MCu32a_R	TAGACTAAAAAGGGTGGATGAG			
Lc_MCu33_F	TCGTTCGCCTCTTCTTCTTC	$(CT)_{24perf}$	222	58
Lc_MCu33_R	GTCATAACCCACGGAGGAGT			
Lc_MCu34a_F	TACAAACCCCAGCCTCTC	$(TC)_4...(TC)_5...(TC)_4...(CT)_5...(TC)_{10}..(CT)_5...(CT)_{4imperf}$	275	48
Lc_MCu34a_R	AGTCTTGATGGAGGTGAGAGT			
Lc_MCu35_F	ACTCTGAATCCTCACCACACA	$(GA)_3C(AG)_6...(AGA)_{5imperf}$	170	58
Lc_MCu35_R	ACATGAGTGCCACACACCAT			
Lc_MCu36a_F	ACATTGCCTCCATGTCGTTA	$(CT)_4TCG(TC)_3G(CT)_{3imperf}$	109	55

Table 1 (continued)

Primers	5′→3′	Microsatellite motifs	Expected size (bp)	Temperature (°C)
Lc_MCu36a_R	TAAGGGTTTTGTTGGGATCG			
Lc_MCu37_F	TCTCCCCTTCTTCCATCTCA	$(CT)_{5}...(CT)_{4imperf}$	150	56
Lc_MCu37_R	GTTTAAACGAATTCGCCCTTC			
Lc_MCu38a_F	TCCTCTTCTCTGTTTTCCAA	$(CT)_{4}...(CT)_{3}C(CT)_{3}....(CT)_{3}C(CT)_{2imperf}$	179	48
Lc_MCu38a_R	AGAGCAACGAAGGAGTACC			
Lc_MCu39a_F	GAATGATTGGGTAACTGCAT	$(TG)_{4}(CA)_{7}(CT)_{7comp}$	243	50
Lc_MCu39b_R	TCACCAATATGATTTTTCACAC			
Lc_MCu40_F	GTCCTCACTCTCTCGCGTTC	$(TC)_{9perf}$	177	55
Lc_MCu40_R	GAAGGAGTATGCAACCCAACA			
Lc_MCu41a_F	TGTGTGAGGAAGATGATGAA	$(GA)_{4}GC(GA)_{10}...(GT)_{6imperf}$	193	48
Lc_MCu41a_R	AAGGAGTTCACACACACACA			
Lc_MCu42_F	CCGTCGGTCGTTTCATTCTA	$(CT)_{6}...(CT)_{4}...(TC)_{5}...(TC)_{4imperf}$	291	55
Lc_MCu42_R	TTTAAACGAATTCGCCCTTC			
Lc_MCu43_F	TCATAAAGCATTTGGCTAAAACA	$(AG)_{9perf}$	154	50
Lc_MCu43_R	CGCAAGCCTCAAGCCTATAA			
Lc_MCu44_F	TGAAGCCTAACCCTAGTATCTCA	$(AC)_{4}...(CA)_{4imperf}$	243	55
Lc_MCu44_R	GGGTTATCTTCCTCGCCTCT			
Lc_MCu45a_F	GCAGCTGAAGAAAAAGAGAA	$(CT)_{5}...(CT)_{5imperf}$	150	60
Lc_MCu45a_R	GTGCTTTAATGGCTTGAGAG			
Lc_MCu46a_F	AGGGTGAAACCCTTATTCTC	$(CT)_{17}...\ (CT)_{4imperf}$	151	48
Lc_MCu46a_R	ATTCTTGAGAAAGGGGAGAG			
Lc_MCu47a_F	TTAGTTCGGAGAGCGTTTAG	$(CT)_{15}...(AC)4\ _{imperf}$	150	48
Lc_MCu47a_R	TGAAGAAGTGGAGGAGAAGA			

Table 1 (continued)

Primers	5′→3′	Microsatellite motifs	Expected size (bp)	Temperature (°C)
Lc_MCu48a_F	TTGAATGATTGAGAGGGTTC	$(GA)_4...(AG)_{4imperf}$	300	55
Lc_MCu48a_R	GACCAGCAAGAATATTGAGC			
Lc_MCu49a_F	ATAAATCCCCATTTCCTTTC	$(TC)_4...(TC)_{5imperf}$	193	55
Lc_MCu49a_R	CTAAAAAGGGTGGATGAGTG			
Lc_MCu50a_F	AACTATGGGGAAGGGAGGAA	$(CT)_7...(TC)_5...(CT)_5...(CT)_{5imperf}$	228	58
Lc_MCu50a_R	ACGAAGGAGTCCACGAGAGA			
Lc_MCu51a_F	CAAACTCAGTGCCGTTACGA	$(TC)_{9perf}$	169	58
Lc_MCu51a_R	GGGAGAGCAACGAAGGAGTA			
Lc_MCu52a_F	TGTGTGAGGAAGATGATGAA	$(GA)_{10}...(GT)_{6imperf}$	149	62
Lc_MCu52a_R	GAGAGCAACGAAGGAGTTC			
Lc_MCu53_F	AGAACCGATCGTGTTTGAGC	$(GA)_5...(GA)_3...(GA)_3...(GA)_3...(GA)_4...(GA)_{5imperf}$	245	55
Lc_MCu53_R	GGGTGGAAAACCCTAATTCC			

Lc *Lens culinaris*, *perf* perfect, *imperf* imperfect, *comp* compound repeats

located within imperfect repeats. Tetranucleotide repeats for primer design were not encountered. Of the 53 SSR primers developed, 71.6% (38 pairs) contained imperfect repeats, 26.4% (14 pairs) contained perfect dinucleotide repeats and 1.8% (1 pair) contained compound repeats. The number of repeat motifs at the perfect AG/AC loci ranged from a minimum repeat length of 5 (Lc_MCu9, Lc_MCu20, Lc_MCu27) to a maximum repeat length of 24 (Lc_MCu33) (Table 1). The duplicated sequences (5 of them) were removed.

Microsatellite Polymorphisms and Genetic Diversity Analyses

All of the developed SSR markers were initially PCR-tested and optimized in Fırat-87, Tigris and Seyran-96 lentil cultivars. In PCR reactions, while Lc_MCu8, Lc_MCu13a, Lc_MCu16b, Lc_MCu29a, Lc_MCu30, Lc_MCu41a, Lc_MCu43, and Lc_MCu44 primers yielded non-specific bands, amplification was not achieved in Lc_MCu37 primer. The remaining 44 SSR primers were analyzed in Applied Biosystems Prism 3500 Genetic Analyzer System (Applied Biosystems, Foster City, CA, USA) and 31 (70.4%) SSR markers were identified as polymorphic and 12 SSR markers (29.6%) were identified as monomorphic for tested cultivars.

The 31 polymorphic SSR markers had 144 alleles in 24 cultivars and the number of alleles per locus varied between 2 and 15 with an average value of 4.64 (Table 2). The Lc_MCu33 primer had the greatest number of alleles (15 alleles). This primer was followed by the primers Lc_MCu19, Lc_MCu24, and Lc_MCu47 respectively with 10, 9, and 9 alleles. Expected heterozygosity ratios of 31 polymorphic SSR markers varied between 0.218 (Lc_MCu31) and 0.903 (Lc_MCu33) with an average value of 0.588 and observed heterozygosity ratios varied between 1.000 (Lc_MCu3, Lc_MCu4, Lc_MCu7, Lc_MCu28, Lc_MCu42) and 0.000 (Lc_MCu50) with an average value of 0.506. Polymorphic information content (PIC) values varied between 0.194 (Lc_MCu31) and 0.895 (Lc_MCu33) with an average value of 0.520.

The dendrogram created with 31 polymorphic SSR markers in 24 cultivars had two different groups (Fig. 1). The first group is composed of Altintoprak, Emre-20, Tigris, Fırat-87, Cagil and Seyran-96 cultivars and the remaining cultivars constituted the second group which was divided into sub-groups. The greatest genetic similarity (91%) was observed between the cultivars Emre-20 and Tigris. The other cultivars with high genetic similarity were identified as Seyran-96 and Cagıl (89%), Bozok and Karagul (88%) cultivars.

Discussion

In the present study, about 350 clones were inquired in Kafkas lentil cultivar from the genomic libraries enriched with AG and AC repeats and 68 (19.4%) of these PCR-screened clones contained SSR motifs. These 68 clones were sequenced and it was observed that 53 (79.1%) sequences contained 134 SSR motifs for primer design. Of the developed SSR markers, 31 (58.4%) were polymorphic. The percentage of SSR motif-containing clones of the present study was 106 times greater than

Table 2 Genetic parameters for SSR primers, number of alleles (n), expected heterozygosity (H_e) and observed heterozygosity (H_o), PIC (polymorphism information content)

Locus	n	H_e	H_o	PIC
Lc_MCu1	2	0.499	0.958	0.374
Lc_MCu2	5	0.691	0.666	0.639
Lc_MCu3	2	0.500	1.000	0.375
Lc_MCu4	4	0.665	1.000	0.603
Lc_MCu5a	8	0.733	0.166	0.707
Lc_MCu7	2	0.500	1.000	0.375
Lc_MCu10	7	0.811	0.333	0.784
Lc_MCu12b	3	0.502	0.833	0.395
Lc_MCu14b	7	0.727	0.333	0.698
Lc_MCu17	2	0.444	0.083	0.345
Lc_MCu18	5	0.615	0.125	0.545
Lc_MCu19	10	0.855	0.333	0.840
Lc_MCu20	3	0.569	0.583	0.504
Lc_MCu21	2	0.304	0.291	0.258
Lc_MCu22	3	0.254	0.291	0.230
Lc_MCu23a	4	0.599	0.500	0.530
Lc_MCu24	9	0.694	0.500	0.664
Lc_MCu28	2	0.500	1.000	0.375
Lc_MCu31	2	0.218	0.250	0.194
Lc_MCu32a	2	0.486	0.166	0.367
Lc_MCu33	15	0.903	0.500	0.895
Lc_MCu34a	6	0.727	0.625	0.692
Lc_MCu35	4	0.553	0.125	0.493
Lc_MCu38a	6	0.730	0.958	0.684
Lc_MCu42	2	0.500	1.000	0.375
Lc_MCu45a	5	0.723	0.500	0.675
Lc_MCu47a	9	0.764	0.250	0.731
Lc_MCu49a	3	0.598	0.125	0.514
Lc_MCu50a	3	0.538	0.000	0.385
Lc_MCu52a	3	0.483	0.750	0.385
Lc_MCu53	4	0.546	0.458	0.494
Total	144	18.231	15.702	16.125
Mean	4.64	0.588	0.506	0.520

the value reported by Hamwieh et al. (2009), but 3.8 and 4.3 times lower than the values reported respectively by Verma et al. (2014) and Andeden et al. (2015) who used enriched method. In present study, 58.4% (31 pairs) of the developed markers were polymorphic. Resultant polymorphic marker percentage was 84% (122 pairs) lower than the value reported by Verma et al. (2014), but 32% (56 pairs) greater than the value reported by Hamwieh et al. (2009) and 23.5% (71 pairs) greater than the value reported by Andeden et al. (2015). Such differences resulted from the methodological approach used in creation of the libraries, selection of restriction enzymes,

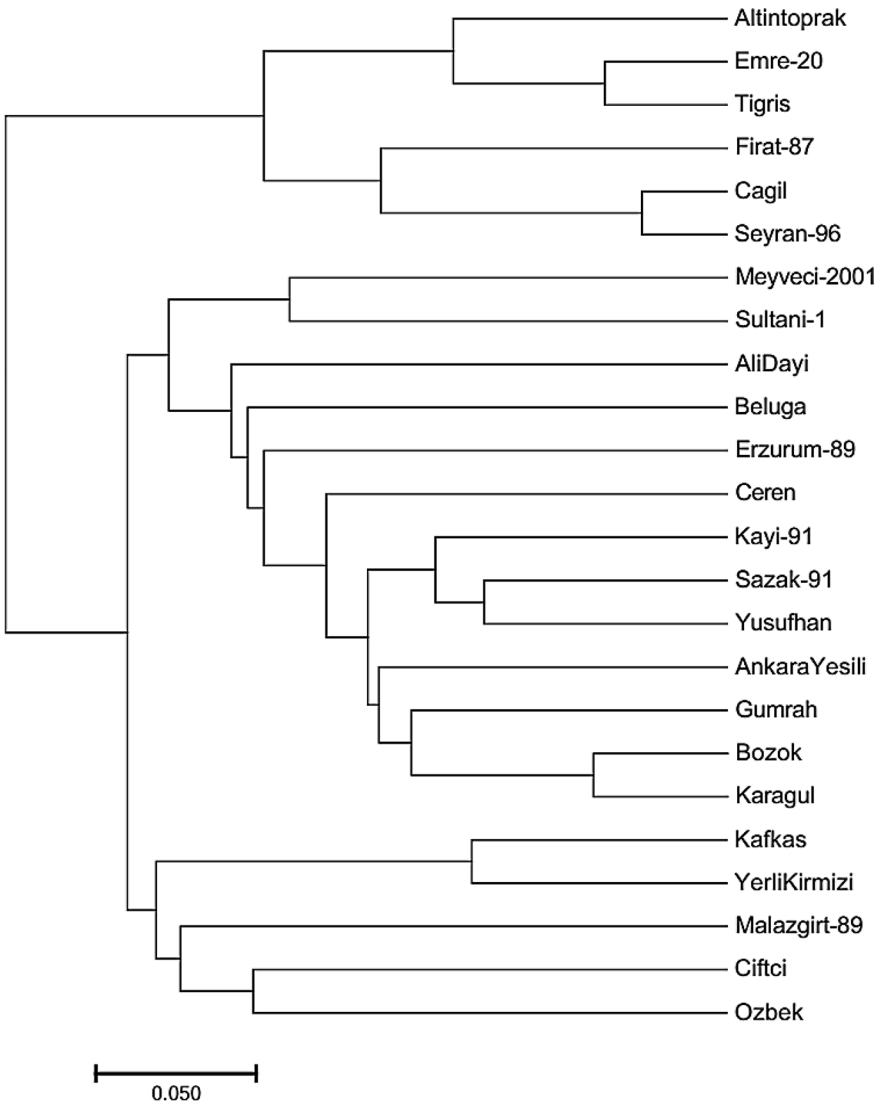

Figure 1 The UPGMA based genetic relationship dendrogram for registered Turkish lentil cultivars formed with the SSR primers developed in this study

types of SSR motifs selected for enrichment or rarity of selected SSR motifs in relevant plant genome (Cuc et al. 2008).

The genetic relationship dendrogram created with 24 registered lentil cultivars used to identify polymorphism ratios of the primers revealed that the developed primers were able to separate all lentil cultivars efficiently. Andeden et al. (2015) also used 8 of present cultivars to test the primers developed and reported similar distribution of these 8 cultivars within the genetic relationship dendrogram.

Greater dinucleotide repeats were reported in lentil genome than the trinucleotide repeats (Hamwieh et al. (2009). Similarly, in the present study, almost all of the SSR primers were composed of dinucleotide repeats and trinucleotide repeats were mostly located in imperfect primer groups. Such a case could be related to the method used. With the method developed by Techen et al. (2010), mostly dinucleotide repeats could be isolated.

Hamwieh et al. (2005) tested SSR primers in *Lens culinaris* sub-species (*L. culinaris* subsp. *culinaris, L. culinaris* subsp. *orientalis, L. culinaris* subsp. *tomentosus, L. culinaris* subsp. *odemensis*) and reported the total number of alleles as 182 with 13 alleles per locus. Total number of alleles for *L. culinaris* subsp. *culinaris* was reported as 128 and number of alleles per locus varied between 2 and 16 with an average value of 9.14. Verma et al. (2014) tested 33 primer pairs in 46 genotypes (*Lens culinaris* sub-species and 8 different legumes) and reported the total number of alleles as 123 and number of alleles per locus as between 2 and 5 with an average value of 3.73. PIC values were reported as between 0.13 and 0.99 with an average value of 0.66. Andeden et al. (2015) tested 78 polymorphic markers in 15 genotypes and reported the total number of alleles as 400 and number of alleles per locus as between 2 and 11 with an average value of 5.1. PIC values were reported as between 0.07 and 0.89 with an average value of 0.58. Present findings on average number of alleles per locus were greater than the values of Verma et al. (2014), similar with the values of Andeden et al. (2015) and lower than the values of Hamwieh et al. (2005). Such differences mostly resulted from differences in number of genotypes and diversity of these genotypes. PIC values of the previous studies and the present study were close to each other.

Up to now, 244 SSR markers were developed for lentils by using genomic libraries (Hamwieh et al. 2005, 2009; Verma et al. 2014; Andeden et al. 2015). With this study, 31 additional new polymorphic SSR markers were developed and the previous number of available SSR markers was raised to 275. These newly developed SSR markers will constitute useful tools for molecular breeding, mapping, assessments of genetic diversity and population structure of lentils.

Acknowledgements This study was supported by TUBITAK (The Scientific and Technical Research Council of Turkey, TOVAG-215O088) and Erciyes University, Scientific Research Projects Unit (FDA-2017–7099). Thanks are extended to Betül-Ziya Eren Genome and Stem Cell Center of Erciyes University for their supports provided throughout the experiments. The authors are also grateful to Assoc. Prof. Dr.Zeki Gökalp (a Certified English Translator and an expert in Biosystems Engineering) for his critical reading and through syntactic corrections of the manuscript.

Funding This study was funded by TUBITAK (Grant Number 215O088) and Erciyes University Scientific Research Projects Coordination Unit (Grant Number FDA-2017–7099).

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Arumuganathan K, Earle ED (1991) Nuclear DNA content of some important plant species. Plant Mol Biol Rep 9:208–218

Andeden EE, Baloch FS, Çakır E, Toklu F, Özkan H (2015) Development, characterization and mapping of microsatellite markers for lentil (*Lens culinaris*Medik.). Plant Breed 134:589–598

Alghamdi SS, Khan AM, Ammar MH, El-Harty EH, Migdadi HM, Abd El-Khalik SM, Al-Shameri AM, Javed MM, Al-Faifi SA (2014) Phenological, nutritional and molecular diversity assessment among 35 introduced lentil (*Lens culinaris* Medik.) genotypes grown in Saudi Arabia. Int J Mol Sci 15:277–295

Ates D, Aldemir S, Alsaleh A, Erdogmus S, Nemli S, Kahriman A, Ozkan H, Vandenberg A, Tanyolac B (2018) A consensus linkage map of lentil based on DArT markers from three RIL mapping populations. PLoS ONE 13(1):e0191375. https://doi.org/10.1371/journal.pone.0191375

Bloor PA, Barker FS, Watts PC, Noyes HA, Kemp SJ (2001) Microsatellite libraries by enrichment. https://www.genomics.liv.ac.uk/animal/RESEARCH/MICROSAT.PDF Accessed 10 April 2018

Cubero JI (1981) Taxonomy, distribution and evolution of the lentil and its wild relatives. In: Witcombe J, Erskine W (eds) Assessment of genetic diversity in lentils (Lens culinaris Medik) based on SNPs. M. Nijhoff & W. Junk Publisher, Boston, pp 187–204

Cuc LM, Mace ES, Crouch JH, Quang VD, Long TD, Varshney RK (2008) Isolation and characterization of novel microsatellite markers and their application for diversity assessment in cultivated groundnut (*Arachis hypogaea*). BMC Plant Biol 8:55

Dikshit HK, Singh A, Singh D, Aski MS, Prakash P, Jain N, Meena S, Kumar S, Sarker A (2015) Genetic diversity in Lens species revealed by EST and genomic simple sequence repeat analysis. PLoS ONE 10:0138101. https://doi.org/10.1371/journal.pone.0138101

El-Nahas A, El-Shazly H, Ahmed S, Omran A (2011) Molecular and biochemical markers in some lentil (*Lens culinaris* Medik.) genotypes. Ann Agric Sci 56:105–112

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

FAOSTAT (2013) https://faostat3.fao.org/faostat-gateway/go/to/download/Q/QC/E. Accessed 7 April 2018

Ferguson ME, Robertson LD, Ford-Lloyd BV, Newbury HJ, Maxted N (1998) Contrasting genetic variation amongst lentil landraces from different geographical origins. Euphytica 102:265–273

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. Nucleic Acids Symp Ser 41:95–98

Havey MJ, Muehlbauer FJ (1989) Variability for restriction fragment lengths and phylogenies in lentil. Theor Appl Genet 77:839–843

Hendre PS, Aggarwal RK (2007) DNA Markers: development and application for genetic improvement of coffee. In: Varshney RK, Tuberosa R (eds) Genomics-assisted crop improvement. Springer, Dordrecht, pp 399–434

Huang X, Madan A (1999) CAP3: A DNA sequence assembly program. Genome Resources 9:868–877

Hamwieh A, Udupa SM, Sarkar A, Jung C, Baum M (2009) Development of new microsatellite markers and their application in the analysis of genetic diversity in lentils. Breed Sci 59:77–86

Hamwieh A, Udupa SM, Choumane W, Sarker A, Dreyer F, Jung C, Baum M (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

Idrissi O, Udupa SM, Houasli C, De Keyser E, Van Damme P, De Riek J (2015) Genetic diversity analysis of Moroccan lentil (*Lens culinaris* Medik.) landraces using simple sequence repeat and amplified fragment length polymorphisms reveals functional adaptation towards agro-environmental origins. Plant Breed 134:322–332

Koressaar T, Remm M (2007) Enhancements and modifications of primer design program Primer3. Bioinformatics 2310:1289–1291

Kahraman A, Kusmenoglu I, Ayidn N, Aydogan A, Erskin W, Muehlbauer FJ (2004) QTL mapping of winter hardiness genes in lentil. Crop Sci 44:13–22

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

Liu K, Muse SV (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics 21:2128–2129

Lefort F, Lally M, Thompson D, Douglas GC (1998) Morphological traits, microsatellite fingerprinting and genetic relatedness of a stand of elite oaks (*Q. robur* L.) at Tullynally Ireland. Silvae Genet 47:5–6

Nei M (1973) Analysis of gene diversity in subdivided populations. Proc Natl Acad Sci 70:3321–3323

Powell W, Mackray GC, Provan J (1996) Polymorphism revealed by simple sequence repeats. Trends in Plant Sci 1:215–222

Phan HTT, Ellwood SR, Hane JK, Ford R, Materne M, Oliver RP (2007) Extensive macrosynteny between Medicago truncatula and *Lens culinaris* ssp. *culinaris*. Theo Appl Genet 114:549–558

Schuelke M (2000) An economic method for the fluorescent labelling of PCR fragments. Nat Biotechnol 18:233–234

Sonnante G, Pignone D (2001) Assessment of genetic variation in a collection of lentil using molecular tools. Euphytica 120:301–307

Sharma SK, Knox MR, Ellis THN (1996) AFLP analysis of the diversity and phylogeny of *Lens* and its comparison with RAPD analysis. Theor Appl Genet 93:751–758

Saha GC, Sarker A, Chen W, Vandemark GJ, Muehlbauer FJ (2013) Inheritance and linkage map positions of genesconferring agromorphological traits in Lens culinaris Medik. Int J Agron 618926 9:pages. https://doi.org/10.1155/2013/618926

Tsanakas GF, Mylona PV, Koura K, Polidoros AN, Gleridou A (2018) Genetic diversity analysis of the Greek lentil (Lens culinaris) landrace 'Eglouvis' using morphological and molecular markers. Plant Genet Resour. https://doi.org/10.1017/S147926211800009610.1017/S1479262118000096

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 24:1596–1599

Tanyolac B, Ozatay S, Kahraman A, Muehlbauer F (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

Techen N, Arias RS, Glynn NC, Pan Z, Khan IA, Scheffler BE (2010) Optimized construction of microsatellite-enriched libraries. Mol Ecol Resour 10:508–515

Temnykh S, DeClerck G, Lukashova A, Lipovich L, Cartinhour S, McCouch SR (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, Karaköy T, Hakle I, Bicer T, Brandolini A, Kilian B, Ozkan H (2009) Genetic variation among lentil (*Lens culinaris* Medik.) landraces from Southeast Turkey. Plant Breed. 128:178–186

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG (2012) Primer3—new capabilities and interfaces. Nucleic Acids Res 40(15):e115

Verma P, Sharma TR, Srivastava PS, Abdin MZ, Bhatia S (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

You FM, Huo N, Gu YQ, Luo MC, Ma Y, Hane D, Lazo GR, Dvorak J, Anderson OD (2008) BatchPrimer3: a high throughput web application for PCR and sequencing primer designi. BMC Bioinform 9:253. https://doi.org/10.1186/1471-2105-9-253