TÜBİTAK

Turkish Journal of Agriculture and Forestry

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

Research Article

Turk J Agric For
(2017) 41: 305-315
© TÜBİTAK
doi:10.3906/tar-1705-105

De novo transcriptome assembly and SSR marker development in apricot (*Prunus armeniaca*)

Muhammed Ali KÖSE[1,2], Necati ÇETİNSAĞ[1,2], Kahraman GÜRCAN[1,2,*]
[1]Genome and Stem Cell Research Center, Erciyes University, Kayseri, Turkey
[2]Department of Agricultural Biotechnology, Erciyes University, Kayseri, Turkey

Received: 25.05.2017 • **Accepted/Published Online:** 10.07.2017 • **Final Version:** 25.08.2017

Abstract: Apricot (*Prunus armeniaca*) is an important fruit crop worldwide. We have performed a de novo transcriptome assembly for 7 apricot accessions ('Stark Early Orange' (SEO), 'Hacıhaliloğlu' (HH), 'Perfection', 'Iğdır', 'Roxana', 'Esen1', and 'Esen2'), which yielded a total number of transcripts ranging from 30,363 for 'SEO' to 59,751 for 'Iğdır'. The pool of the reads produced from 7 accessions were assembled into 85,766 transcripts, with an average of 1165.69 nt. Functional annotation (Gene Ontology- GO and Kyoto Encyclopedia of Genes and Genomes- KEGG) was performed successfully for the transcripts. Simple sequence repeats (SSRs) were searched in the transcript pool and 14,722 di-, tri- tetra-, penta-, and hexanucleotide motif loci with a minimum of 5 repetitions for all motifs were identified. Primers were designed for 206 loci, and 72 of them were found to be polymorphic by amplifying diverse 24 apricot accessions, including 7 *Plum Pox Virus* (PPV)-resistant and 17 PPV-susceptible accessions. In order to test the amplification success of publicly available genomic SSRs (gSSRs) for diverse apricot accessions, an additional 88 published *Prunus* gSSRs were characterized amplifying the same 24 apricots and only 54 (62%) produced polymorphic bands. The new EST-SSRs could be a reliable source of primers for characterization and mapping studies of apricots, especially because they mostly flank easily scorable tri- and tetranucleotide repeats.

Key words: Microsatellite, molecular markers, functional annotation

1. Introduction

The common apricot (*Prunus armeniaca* L.) is diploid, with eight pairs of chromosomes (2n = 16) and an estimated genome size of 240 million nucleotides (nt) according to the Genome Database for Rosaceae (https://www.rosaceae.org/). Apricot belongs to the family Rosaceae, subfamily Prunoideae, and it is native to China and Central Asia, which are two primary genetic diversity centers for the species. The near-eastern group, including Turkey, Iran, and the Caucasus, is considered as a secondary center of diversity (Vavilov, 1951; Ercisli, 2009; Halasz et al., 2010; Hegedus et al., 2010).

High-throughput next-generation sequencing (NGS) technologies produce large amounts of data and are thus widely used for transcriptome analysis, allowing quantification of RNA transcripts, discovery of new genes, and a vast amount of polymorphic loci. The potential of NGS in apricot science was reviewed by Martínez-Gómez et al. (2011). NGS technologies have been applied in apricot species to facilitate the transcriptome analysis of several biological and agronomical aspects: seasonal bud dormancy (Zhong et al., 2013) and self- and cross-pollinated pistils (Habu et al., 2014) in Japanese apricot, global gene profiling and the search for potential SSR markers (Dong et al., 2014), oil dynamic accumulation in developing seed kernels for the development of woody biodiesel (Niu et al., 2015) in Siberian apricot (*Prunus sibirica* L.), *Plum Pox Virus* (PPV) (Sharka) susceptibility/resistance (Rubio et al., 2015), single nucleotide polymorphism (SNP) discovery (Salazar et al., 2015), the study of the development of embryos (Bai et al., 2016), and SNP discovery and genetic characterization via genotyping by sequencing in common apricot (Gürcan et al., 2016).

Simple sequence repeats (SSRs) have been highly preferred due to their high variability, codominant inheritance, suitability for sharing among laboratories, and cross-species transferability (İpek et al., 2016; Sakar and Ünver, 2016; Sorkheh and Khaleghi, 2016). A variety of SSR markers have been developed for *Prunus*. Examples include almond (*Prunus dulcis*) (Testolin et al., 2004; Messina et al., 2004), apricot and Japanese apricot (*Prunus mume*) (Lopes et al., 2002; Decroocq et al., 2003; Vilanova et al., 2006; Li et al., 2010; Wang et al., 2014), cherries (Clarke and Tobutt, 2003; Sorkheh et al., 2016),

* Correspondence: kgurcan@erciyes.edu.tr

and peach (Aranzana et al., 2002; Yamamoto et al., 2002; Howad et al., 2005; Chen et al., 2014; Dettori et al., 2015). SSR markers along with other marker systems have been used in the construction of apricot genetic maps, genetic diversity assessments, and characterization of apricot germplasm collections (Lambert et al., 2007; Lalli et al., 2008; Dondini et al., 2011; Soriano et al., 2012; Rubio et al., 2014, Decroocq et al., 2014, 2016; Gürcan et al., 2015).

Although a variety of SSRs have been identified in *Prunus* species including apricot, more reliable SSR markers are needed for apricot molecular breeding, particularly for anchoring the parental maps and construction of saturated maps of apricot. Available primers occasionally fail during polymerase chain reaction (PCR), probably due to their having not been tested on a wide set of apricots representing overall apricot genetic diversity. Furthermore, the useful number of SSRs lowers drastically while mapping them to biparental segregation populations since SSRs do not always produce polymorphic alleles for parental accessions. Additionally, the majority of available primers were usually developed from SSR-enriched genomic libraries, derived primarily from intergenic DNA regions. In contrast, EST-SSRs are specifically developed from transcribed regions of the genome and present high potential for linkage to loci of interest. Thus, polymorphic EST-SSRs are valuable in constructing linkage maps, presenting considerable utility for MAS. Here, we report the NGS sequencing and transcriptome profiling of 7 apricot accessions and the development of 72 polymorphic EST-SSR loci obtaining allele sizes of 24 diverse accessions, including 7 PPV-resistant and 16 PPV-susceptible accessions. Additionally, we have studied 88 previously published gSSRs amplifying the same diverse 24 apricots in order to exhibit how many of the published *Prunus* primers are useable in apricot breeding programs.

2. Materials and methods

2.1. Plant material

We used 7 accessions including international, national, and local accessions for the transcriptome analysis: 'SEO', 'Hacıhaliloğlu' (HH), 'Perfection', 'Iğdır', 'Roxana', 'Esen1', and 'Esen2'. 'SEO' is PPV-resistant and the most commonly used donor for PPV resistance breeding programs in Europe. 'HH' is susceptible to PPV and accounts for most dried apricots (about 70%) in Turkey, and thereby is also the main cultivar for apricot studies and breeding programs in Turkey. 'Iğdır', 'Perfection', and 'Roxana' are important apricots for fresh consumption in Turkey, the United States, and Europe, respectively. 'Esen1' and 'Esen2' are local accessions preferred in the Kayseri region and are highly susceptible to PPV. For characterization of SSR loci, 24 accessions representing genetic diversity in apricot were used including 7 PPV-resistant ('SEO', 'Harlayne', 'Zard', 'M2243', 'M2244', '2_7', 'No 8 Zerdali') and 17 susceptible apricots ('1297', '2254', 'Adilcevaz 2', 'Gü 2', 'Güz Eriği', 'Hacıhaliloğlu', 'Harcot', 'Levent', 'Markuleşti', 'Mektep', 'No 8 Mektep', 'Polumella', 'Roxana', 'Şekerpare Benzeri', 'Y5', 'Y6')

2.2. RNA isolation, library construction, and Illumina sequencing

Fresh leaf tissue (100 mg) was used for total RNA extraction via a commercial plant RNA isolation kit. Integrity, quantity, and quality of the RNA of the 7 samples were determined using a spectrophotometer and a 2100 Bioanalyzer RNA nano chip device (Agilent, Santa Clara, CA, USA). RNA samples were used for sequencing when the RNA Integrity Number (RIN) was >7. The cDNA libraries were constructed and sequenced using an Illumina HiSeq 2000, located at the Genetic Engineering and Biotechnology Institute (GEBI) of TÜBİTAK, Turkey.

2.3. De novo transcriptome assembly and functional and structural annotation

Initially, clean reads were obtained using a FASTX-Toolkit by removing adaptor sequences, discarding the low-quality reads (Phred quality score less than 20, $Q \geq 20$) and ambiguous nucleotides ('N') at the end of reads, and finally removing the short reads (length below 50 nt). The high-quality reads were then assembled to reconstruct transcripts for each sample using Trinity software (Haas et al., 2013). All reads of seven accessions were also combined and used to assembly transcripts, which is called the 'pool' of transcripts. Functional annotation (Gene Ontology-GO and Kyoto Encyclopedia of Genes and Genomes-KEGG) was performed by means of sequence similarity (e-value cutoff of 10^{-5}) using Trapid (Van Bel et al., 2013), and SSRs were identified using GMATo (Wang et al., 2013) and SSRIT (http://www.gramene.org/db/markers/ssrtool). Transcripts were aligned to Japanese apricot (*Prunus mume*) and peach (*Prunus persica*) genomes using Blast2go software (Conesa et al., 2005).

2.4. Development and characterization of EST-SSR primers and assessment of published gSSR primers

After determining SSR loci in transcripts, primers were designed for 206 randomly selected loci with long internal repeats (≥20 nt) and flanking sequences sufficient for primer design. The PRIMER3 program (Rozen et al., 2000) was used for designing primers with the following criteria: 22–24 bp length, 30%–55% GC content, and 60 °C annealing temperature. Initial screening of markers in order to identify polymorphic loci was performed by amplifying DNA of 24 apricot genotypes and running agarose gels. The PCRs were conducted in a volume of 15 µL containing water, 12 pM each of forward and reverse

primers, 1X reaction buffer, 35 µM $MgCl_2$, 27 µM each of dNTPs, ~25 ng of template DNA, and 0.25 U of DNA polymerase. The thermal cycler was programmed for denaturation at 94 °C for 5 min followed by 35 cycles of 94 °C for 40 s, 60 °C for 30 s for primer annealing, and 72 °C for 40 s with a final 7-min extension step at 72 °C. The products were run on 2% agarose gels. Fluorescently labeled SSR genotyping and PCR was performed according to the method described by Schuelke (2000) for allele sizing and characterization of the loci that appeared polymorphic in the initial screening on agarose gels. PCR reactions were performed including M13 primer fluorescently labeled with 6-FAM, NED, PET, or VIC amplifying the same 24 apricots. For multiplexing, 1 µL of labeled products of each of four primers was combined with distilled water to a final volume of 200 µL. A 1-µL aliquot was loaded into an ABI 3500 capillary electrophoresis instrument (Applied Biosystems, Foster City, CA, USA). For marker characterization, a genetic similarity matrix based on the proportion of shared alleles was generated and the expected heterozygosity (He), observed heterozygosity (Ho), and polymorphism information content (PIC) were calculated using PowerMarker V3.025 software (Liu and Muse, 2005). In order to increase reliable SSR loci and moreover to exhibit how many of the existing SSRs are indeed usable in apricot breeding programs, a total of 88 previously published gSSRs were also amplified using the same 24 accessions. They included 20 primers developed in almond (Testolin et al., 2004), 50 in apricot (Lopes et al., 2002; Messina et al., 2004; Vilanova et al., 2006), 8 in tetraploid cherry (Cantini et al., 2001), and 10 in peach (Sosinski et al., 2000).

3. Results

3.1. De novo transcriptome assembly, functional annotation, and classification

After removing low-quality sequences and trimming adapter and ambiguous nucleotides ('N') at the end of reads, de novo assembly of all clean reads was obtained for the seven accessions (Table 1). De novo assembly produced an average of 43,023 sequences with an average N50 of 1762 nt. Among the seven accessions, apricot 'Iğdır' produced the highest number of transcripts (59,751), and total length (76,625,300 nt), as well as average length (1282.41 nt) and maximum length (15,046 nt). The average GC content ranged from 41.75% ('Iğdır') to 43.67% ('SEO'). After combining the quality reads from all 7 samples, the assembly of reads generated a total of 85,766 pool transcripts. The maximum and average length of the pool transcripts were 16,371 and 1165, respectively. Among the pool transcripts, the majority (35,306) ranged between 201 and 500 nt, followed by 15,163 transcripts in the size range of 501–1000 nt. The number of transcripts for the pool with a size over 3000 nt became 6359 and N50 was 2006 nt.

We performed annotation for the transcripts of 7 accessions and the pool. GO assignments were used to classify the functions of the sequences. The transcripts of the pool (85,766 fragments) were assigned to 158,816 GO terms as depicted in the Figure. Of these GO terms, 65,119 (41%), 46,225 (29.11%), and 47,477 (29.89%) were classified into the three GO main categories of "biological process", "cellular component", and "molecular function", respectively. The transcripts were assigned to 20 functional groups in the biological process category and "metabolic

Table 1. Summary of outputs of sequence analysis.

Sample	Total number of sequences	Total length (nt)	Average sequence length (nt)	Maximum sequence length (nt)	N25	N50	N75	Total GC count (nt)	GC (%)
SEO	30,363	28,532,152	939.70	11,477	2325	1529	818	12,461,102	43.67
Esen1	41,385	35,313,704	853.30	11,364	2058	1331	713	15,297,593	43.32
Esen2	31,002	35,721,316	1152.23	11,464	2626	1777	1090	15,531,837	43.48
Perfection	53,641	62,461,631	1164.44	14,376	2911	1937	1132	26,421,993	42.30
Iğdır	59,751	76,625,300	1282.41	15,046	3096	2066	1268	31,992,294	41.75
Roxana	45,247	51,677,243	1142.11	14,899	2860	1871	1099	22,020,860	42.61
HH	39,778	45,176,279	1135.71	12,972	2756	1824	1084	19,421,704	42.99
Average	43,023	47,929,661	1095.7	NA	2661	1762	1029	20,449,626	42.87
Pool	85,766	99,976,418	1165.69	16,371	3077	2006	1149	41,314,448	41.32

NA, Not applicable.

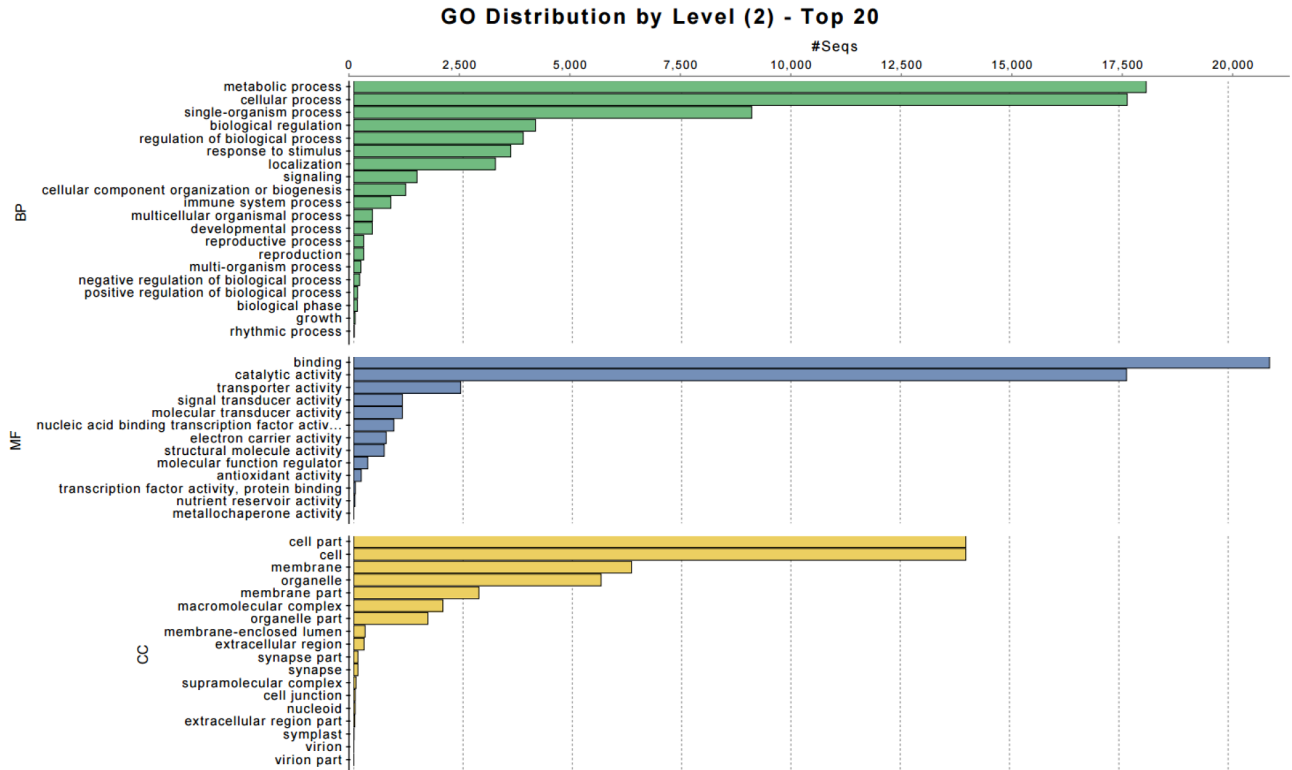

Figure. Gene ontology (GO) distributions of assembled transcripts of the pool.

process" (18,125) and "cellular process" (17,691) dominated, followed by "single-organism process" (9104). With respect to molecular function, 13 functional groups were detected and the "binding" term was the most highly represented term (20,974 transcripts), followed by "catalytic activity" (17,677) and "transporter activity" (2448) processes. Coming to the cellular component category, "cell part" and "cell" were the major terms and both were represented by 14,001 transcripts, followed by the term "organelle" (6355) among the 16 functional groups. With respect to KEGG analysis, a total of 5612, 4657, 5865, 5885, 5848, 5847, and 5923 transcripts of 'SEO', 'Esen1', 'Esen2', 'Perfection', 'Iğdır', 'Roxana, and 'HH' were assigned with 324, 224, 325, 327, 323, 326, and 323 KEGG pathways, respectively.

3.2. SSR mining and characterization of EST-SSR and published gSSRs

To identify the repeat loci proper for SSR marker development, we screened the transcripts of 7 accessions and the pool for di- to hexanucleotide repeat loci with a minimum of 5 repetitions for all motifs. The total number of repeats discovered for each accession ranged from 4356 ('SEO') to 11,037 ('Iğdır') and was 14,722 for the pool (Table 2). Frequencies of SSR motifs for each accession were

Table 2. Distribution to different repeat type classes.

Motif	SEO	Esen1	Esen2	Perfection	Iğdır	Roxana	HH	Pool
Di	2664	3191	3646	5946	7405	4898	4484	10,123
Tri	1616	1814	2027	2866	3386	2367	2420	4263
Tetra	56	76	84	172	187	120	105	288
Penta	11	14	26	31	41	25	26	30
Hexa	9	12	9	25	18	11	12	18
Total	4356	5107	5792	9040	11,037	7421	7047	14,722

determined. For the pool sequences, the most abundant repeat motif was dinucleotide at 10,123 (68.8%), followed by trinucleotide at 4263 (28.95%), tetranucleotide at 288 (1.95%), pentanucleotide at 30 (0.2%), and hexanucleotide at 18 (0.12%). The dinucleotide AG/CT motifs constituted more than three-quarters (8042, 79.44%) of the total number of dinucleotides, followed by AT/AT (1073, 10.6%), AC/GT (1002, 9.9%), and CG/CG (6, 0.059%). Considering trinucleotides, AGG/CTT was the most abundant repeat motif.

Primers were designed for 206 SSR repeats of ≥20 nt of the pool transcripts. In the initial screening, 171 of the 206 primer pairs produced bands, while the remaining 35 failed to produce PCR products. Of these 171, 65 primer pairs produced nonspecific amplification far from SSR loci patterns and 34 produced monomorphic bands and were not investigated further; the remaining 72 generated one or two polymorphic bands in a diploid manner per accession and could be scored with confidence. Of the 72 SSR loci, 33 contained trinucleotide motifs, while 27 and 12 were tetranucleotide and dinucleotide motifs, respectively. The primers AK1, AK164, and AK200 produced alleles 100–200 nt longer than the original locus size determined by the Primer3 program, suggesting the presence of introns on these loci. We obtained PCR products for all accessions at all loci. For the 72 loci, allele sizes were obtained for the 24 accessions in accordance with the SSR locus evolution. A total of 293 alleles were produced at the 72 loci. The number of alleles per locus ranged from 2 to 12, with a mean of 4.07. Mean He, Ho, and PIC were 0.48, 0.66, and 0.43, respectively. PIC values were >0.50 for 25 loci and ranged from 0.50 and 0.90 for these 25 loci (Table 3).

Table 3. Primer sequences and characteristics of new EST-SSR loci for apricot.

Primer sequence (5’-3’)	Size	Motif	n	He	Ho	PIC	Primer sequence (5’-3’)	Size	Motif	n	He	Ho	PIC
ak1f:acccttgcagcttgaaattg	295	aagc	3	0.4	0.4	0.4	ak105f:caaggggatttactcgaagc	156	aca	3	0.5	0.8	0.4
ak1r:ttatctcactcgctgccatc							ak105r:ccaccagctttcttcttgct						
ak3f:ggttggagggtgcataagaa	151	ca	6	0.8	0.8	0.7	ak107f:tccggagttgccaagtctta	257	tca	6	0.6	0.9	0.5
ak3r:tccacacaaaatgtgaagat							ak107r:gatgatgacgaggacgatga						
ak5f:ctgcgtccctgaaacattct	152	ta	5	0.7	0.9	0.6	ak109f:catgctgatggtgatcaggt	132	atg	3	0.5	0.3	0.4
ak5r:aactggccaaaccaatccta							ak109r:gagaggcttcaacagcaagg						
ak7f:atttgaaagggcgaacacaa	191	tg	4	0.6	0.7	0.5	ak110f:gctgctgctgttgttgatgt	140	atg	4	0.5	0.3	0.4
ak7r:gttgagggctcgtttgaatc							ak110r:ttcaacagcaagggaaggac						
ak12f:caccacctccttaaaccct	198	aga	5	0.6	0.7	0.5	ak114f:caaacccttttgatcaacgaa	257	cca	2	0.4	0.5	0.3
ak12r:ccgatccgctactttctttg							ak114r:ttgcaggagcttccagactc						
ak18f:caaacgcattcaatccatca	245	ac	8	0.7	0.6	0.7	ak118f:tgtcgaattgggatgatgaa	258	tgct	4	0.3	0.2	0.2
ak18r:tggaagcaaggtaaatgcag							ak118r:ggtaattgggcttagggtgg						
ak21f:aaatctcgatgccaacaacc	242	agg	3	0.2	0.2	0.2	ak122f:cttgcatctcactcccaacc	254	atgt	5	0.2	0.2	0.2
ak21r:ttggcatcctcctctttcaa							ak122r:tcaccttgtcattctcacgtg						
ak23f:tcctcccatcatgagctttc	149	aacc	4	0.4	0.3	0.3	ak123f:cccaggcctcttccttctta	275	cttt	5	0.7	0.7	0.6
ak23r:ggcttggcaatagtggaaga							ak123r:ttttgtgtttttcttccggg						
ak24f:tctgcatgctgctgtttagc	313	cac	3	0.2	0.1	0.2	ak124f:tcttctctgaacgccctcat	146	cct	7	0.8	0.8	0.8
ak24r:tatatctaccggcccattgc							ak124r:gttgtcgtttaggcggagag						
ak30f:aagccgacatattcaccttg	270	tg	4	0.6	0.3	0.6	ak130f:tacgtaaagcgcacccattt	141	ctc	3	0.5	0.3	0.5
ak30r:caacaatctgtacggccaat							ak130r:caaaactctcgcaaaagcct						
ak32f:tgacatgcgcactcttctct	192	agtg	4	0.3	0.2	0.3	ak132f:ccccacaattacattggaag	264	cag	7	0.7	0.7	0.7
ak32r:cagttggtaggccctggtaa							ak132r:catatcatcagctgcgccta						
ak33f:ggaatcacggtggaaagaga	206	tcta	3	0.2	0.0	0.2	ak133f:ccctctctctccgtatgatc	269	tttc	5	0.7	0.5	0.6
ak33r:aagcgaaatagaagagggag							ak133r:tttcccttccctttcttcgt						
ak36f:gctaagtgccaacgatagcc	178	atta	3	0.1	0.1	0.1	ak138f:gcgcacatcatctctgtctt	270	aat	3	0.4	0.5	0.4
ak36r:ttcattcctggagcctcatc							ak138r:ccaacaaacatggcttctga						
ak37f:tggcatgttgggaagaaaat	174	aatg	5	0.6	0.7	0.6	ak140f:acacggtctgtgtgtgtaga	272	ga	6	0.5	0.3	0.4
ak37r:cctgtatgccatttgcagag							ak140r:tatgggatggtaccaaaaacg						
ak38f:ggcacgagaatctccactgt	234	gct	4	0.5	0.7	0.5	ak143f:actgagtcgtcgccatcttt	268	aat	4	0.5	0.7	0.4
ak38r:gccatcctacatcgacaggt							ak143r:tagtggctggctgtaaagcc						

Table 3. (Continued).

Primer sequence (5'-3')	Size	Motif	n	He	Ho	PIC	Primer sequence (5'-3')	Size	Motif	n	He	Ho	PIC
ak39f:gtgcaaattcaagccaaagc	264	agg	3	0.4	0.3	0.4	ak144f:cccattcacacagctaagca	229	agaa	2	0.4	0.4	0.3
ak39r:gtgagtagggacccaccag							ak144r:aaccaagaactggaaaatcgt						
ak44f:aagaagccggttggaatttt	216	ct	7	0.7	0.5	0.7	ak145f:gcgcagagctatcaaactcc	154	ctg	2	0.3	0.3	0.3
ak44r:agatttcgctccaaagctga							ak145r:gatggttggggtgatgagtt						
ak57f:aggggaccaggaaccaataa	216	tttg	7	0.8	0.6	0.8	ak150f:caatgaagactgtccacacc	250	ataa	3	0.5	0.6	0.4
ak57r:tgtcacaagctcacacctca							ak150r:tcagaaaaggaaattcactg						
ak62f:agcctgcctattcactttgg	274	ttta	2	0.1	0.0	0.1	ak151f:ttggtgtcatttgccacttc	256	ttcc	5	0.3	0.4	0.3
ak62r:ctttgatggtgtggggtttt							ak151r:tgtgggtatctctgagcaac						
ak69f:tgaaactgaggacgatgacg	235	gat	3	0.5	0.3	0.4	ak154f:cctctccattgttttgactg	150	cacg	3	0.2	0.2	0.2
ak69r:cgtcttccggatttgcttta							ak154r:cagaaacccagatagacggc						
ak73f:tgggttgtggttggagtttt	164	aaga	7	0.7	0.7	0.7	ak158f:tttcaaagccttccgttgat	267	tgc	2	0.4	0.4	0.3
ak73r:agttcagagcccaaaagcaa							ak158r:tttgccatttccttttcctg						
ak75f:ctcctcccactccttcattg	142	tttc	2	0.5	0.3	0.4	ak161f:ctttgtaccccaccaaatcg	140	cat	6	0.5	0.6	0.5
ak75r:tgccaaaaccaaattgactg							ak161r:gagacgcagcaggaatctg						
ak80f:gcaccaaaagcaaaggaaga	235	aga	4	0.7	0.7	0.6	ak164f:ctcggagtcgtcaaaacaca	356	cttt	2	0.4	0.5	0.3
ak80r:gagctgtttctgcagccttc							ak164r:tccatggtcattgttgacag						
ak82f:gagaatggcagtggaagctc	270	ctg	3	0.6	0.7	0.5	ak166f:gggtaattcatgtgcccaat	170	tct	3	0.4	0.3	0.4
ak82r:aggcagcaacaacaagaaca							ak166r:cccttttgcttcctccttct						
ak83f:tcctcctcttcatcctgctc	159	tcc	2	0.0	0.0	0.0	ak169f:cagagaagtacagttggatt	144	ag	2	0.5	0.9	0.4
ak83r:gtgaaaatggaggtgggttg							ak169r:gcaatgcactgctcaaacac						
ak85f:caaaaacaaaaacccaacagc	159	ag	9	0.8	1.0	0.8	ak172f:gcttcttccggttcgatttc	142	tgc	5	0.7	0.7	0.6
ak85r:ctgcagcaatcgtacagagc							ak172r:tttgcacccatgaactgaag						
ak87f:tgcagctttaatttcatcaag	259	ataa	6	0.7	0.4	0.6	ak173f:gcttcttccggtcgatttc	254	tgc	5	0.5	0.3	0.5
ak87r:ggttgtgctgtttctttcct							ak173r:tttgcacccatgaactgaag						
ak89f:tttggccaccgacttatctc	167	tcta	5	0.6	0.6	0.6	ak176f:cattggactgtcatcgctgt	134	gaa	5	0.6	0.6	0.5
ak89r:cagtttgctgagcagtcagg							ak176r:gaaacccatgctttctccaa						
ak90f:aatggtgttgttgacgagca	145	tat	5	0.7	0.6	0.6	ak179f:cagtcaatctcggaccgttt	166	agg	2	0.5	1.0	0.4
ak90r:atcacggagcagaagttggt							ak179r:ttcttgcagacgcactcttc						
ak92f:gccaaccaacaaacttgagg	275	ttat	3	0.3	0.1	0.3	ak180f:cccaaggggaatgcaaata	269	tg	4	0.5	0.5	0.4
ak92r:aatcatgccttatatttagac							ak180r:cctatctactgggcactgcc						
ak94f:gcctggggtttgatcaatta	158	gcc	3	0.2	0.2	0.2	ak181f:aacatgggtggatgatgctt	148	ttgc	3	0.2	0.2	0.2
ak94r:ctcttcttggggctctcctt							ak181r:tttgtatgcagctaggcacg						
ak98f:tggaaaaacccaaaagccta	142	gagg	3	0.6	0.4	0.5	ak187f:ccactggtattggtattggta	145	tga	3	0.4	0.4	0.4
ak98r:catgcagcccaactcctaat							ak187r:gtcgagcactctgaaaattta						
ak99f:ctgcaatgccaacaagagaa	258	cagt	2	0.2	0.3	0.2	ak190f:ctgttgttgcatctgaggat	142	tgt	3	0.5	0.1	0.5
ak99r:tctcagcagggtgaaatgaa							ak190r:cacctacttgtggcttaaac						
ak102f:aggaggaagaggccattgt	155	taaa	2	0.2	0.3	0.2	ak193f:gcaaatcagctagtgaaaga	155	ac	12	0.9	0.8	0.9
ak102r:tacaactcctcccacccaac							ak193r:taccactttacgatgtcgtt						
ak103f:ctcctccagccttcttctga	254	gca	3	0.6	0.2	0.5	ak198f:atctacacaccagcagagag	343	catc	3	0.2	0.2	0.2
ak103r:tgctattgcgacgatgatgt							ak198r:ctataatgctctcgcctgag						
ak104f:atagggttgcttggcatcat	236	at	4	0.6	0.4	0.5	ak200f:ccagtagattgggtgctact	426	tca	2	0.5	1.0	0.4
ak104r:cccttgcaagaaaaaccatt							ak200r:ctacgtccaagaacaagatt						
								average			0.5	0.5	0.4

n: Number of alleles, He: expected heterozygosity, Ho: observed heterozygosity, PIC: polymorphism information content.

For the assessment analysis of published SSRs for 24 diverse apricots, 88 primers developed in *Prunus* species were investigated. Of the total 88 primer pairs, 54 produced nice polymorphic bands in apricot and they were further characterized, obtaining allele sizes of 24 apricot accessions. Of the remaining, 15 (17%), and 14 (16%) produced monomorphic and nonspecific amplicons, respectively, while 5 (6%) failed to produce amplicons (Table 4). The 54 polymorphic loci produced 411 alleles and the number of alleles per locus ranged from 2 to 18, with a mean of 7.47. Mean values for He, Ho, and PIC were 0.69, 0.60, and 0.65, respectively (Table 5). PIC values were <0.50 for 8 loci and ranged from 0.50 and 0.90 for the remaining loci, which indicates that they are highly informative.

4. Discussion

In the last decade, high-throughput sequencing technologies together with an increasing number of reliable bioinformatics tools have accelerated genomic and transcriptomic studies in crop breeding. Transcriptome sequencing has become an efficient method to generate high levels of genomic data, large EST sequences, and molecular markers. In this study, we have performed transcriptome analysis of leaves from 7 apricot accessions and obtained an average of 43,023 transcripts (N50: 1762 nt) for all accessions with a mean length 1095 nt. The results are compatible with recently published transcriptome analysis of apricots: Bai et al. (2016) obtained 59,851, 57,163, and 54,792 contigs from three embryo samples of apricot with average contigs size exceeding 427 nt (N50 of 908 nt). Niu et al. (2015) generated 124,070 contigs (N50: 1603 bp) with the mean length of 829.62 bp in Siberian apricot. Habu et al. (2014) obtained 40,061 contigs ranging from 201 to 7477 bases with an average length of 572 bases by using another type of high-throughput technology, 454 pyrosequencing.

In order to perform functional annotation and classification, the assembled transcripts were subjected to GO annotations and KEGG analysis. The transcripts of the pool (85,766 fragments) were classified into the three GO main categories of biological process, cellular component, and molecular function with numbers 65,119 (41%), 46,225 (29.11%), and 47,477 (29.89%), respectively. The present findings are in agreement with previous results; Dong et al. (2014) reported the highest assignments of transcripts of embryos of Siberian apricot to biological processes (54,667, 40.72%), followed by cellular components (51,551, 38.40%) and molecular functions (28,039, 20.88%). Similarly, GO assignment of transcripts sequenced from buds, leaves, stems, flowers, fruit pulp, and seeds of commercial apricot resulted in maximum hits to biological processes, followed by cellular components and molecular function (Bai et al., 2016). The highest hits were related to the biological process GO category, which may indicate that the analyzed tissues were undergoing extensive metabolic activities. Consistent with that, in the present study and in the other two studies "metabolic processes" and "cellular process" were the most frequent terms in the biological GO category; moreover, KEGG analysis revealed that the highest numbers of transcripts were involved in metabolic pathways. Indeed, we used the freshest growing leaves of shoot tips, which explains the extensive metabolic activities found. The predicted GO terms, together with the KEGG analysis, are useful for future research into gene functions.

Here we identified 14,722 di-, tri-, tetra-, penta-, and hexanucleotide motif loci with a minimum of 5 repetitions for all motifs. The dinucleotide AG/CT motifs constituted more than three-quarters (8042, 79.44%), and CG/CG was

Table 4. Amplification and polymorphism of published gSSRs in 24 diverse apricots.

Source plant	Source article	n	Polymorphic loci (%)	Monomorphic loci (%)	No amplification	Nonspecific amplification (%)
Almond	1	20	9 (45)	2 (10)	3 (15)	6 (30)
Apricot	2	22	12 (55)	8 (36)	2 (9)	-
Apricot	3	20	13 (65)	1 (5)	-	6 (30)
Apricot	4	8	5 (63)	1 (13)	-	2 (25)
Cherry*	5	8	6 (75)	2 (25)	-	-
Peach	6	10	9 (90)		-	1 (10)
	Total	88	54(62)	14 (16)	5 (6)	15 (17)

*Tetraploid; n, number of SSRs. 1, Testolin et al. (2004); 2, Lopes et al. (2002); 3, Messina et al. (2004); 4, Vilanova et al. (2006); 5, Cantini et al. (2001); 6, Sosinski et al. (2000).

Table 5. Characteristics of published gSSR loci in 24 apricot accessions.

Accession no.	Source species	Size range (nt)	n	He	Ho	PIC	Accession no.	Source species	Size range (nt)	n	He	Ho	PIC
ssrPaCITA2	Apricot	240–270	8	0.8	0.1	0.7	UDA-002	Almond	148–184	9	0.9	0.3	0.8
ssrPaCITA6	Apricot	154–238	6	0.6	0.0	0.5	UDA-004	Almond	172–174	2	0.3	0.0	0.2
ssrPaCITA7	Apricot	204–232	9	0.7	0.5	0.7	UDA-005	Almond	166–192	7	0.6	0.3	0.6
ssrPaCITA10	Apricot	166–192	10	0.8	0.5	0.8	UDA-013	Almond	187–195	5	0.6	0.2	0.5
ssrPaCITA11	Apricot	143–157	5	0.4	0.3	0.4	UDA-015	Almond	118–142	10	0.8	0.9	0.7
ssrPaCITA12	Apricot	154–174	9	0.8	0.7	0.7	UDA-020	Almond	154–188	11	0.9	1.0	0.9
ssrPaCITA14A	Apricot	158–202	8	0.8	0.9	0.7	UDA-021	Almond	107–137	11	0.8	0.9	0.8
ssrPaCITA18	Apricot	142–180	7	0.8	0.9	0.8	UDA-022	Almond	178–190	5	0.4	0.3	0.4
ssrPaCITA20	Apricot	216–219	2	0.1	0.1	0.1	UDA-025	Almond	130–136	4	0.7	0.9	0.6
ssrPaCITA22	Apricot	182–185	2	0.1	0.1	0.1	aprigms3	Apricot	268–280	7	0.8	0.4	0.8
ssrPaCITA23	Apricot	157–173	7	0.8	0.5	0.8	aprigms8	Apricot	191–223	9	0.8	0.3	0.8
ssrPaCITA27	Apricot	240–280	7	0.8	0.6	0.8	aprigms10	Apricot	292–304	8	0.8	0.7	0.8
PMS2	Cherry*	187–227	7	0.8	0.7	0.8	aprigms11	Apricot	178–184	4	0.7	0.8	0.6
PMS30	Cherry	110–170	10	0.8	0.8	0.8	aprigms16	Apricot	225–257	9	0.8	0.9	0.8
PceGA25	Cherry	183–207	6	0.7	0.5	0.7	UDAp-406	Apricot	125–161	11	0.6	0.8	0.6
PMS40	Cherry	118–132	3	0.4	0.0	0.3	UDAp-407	Apricot	191–223	6	0.7	1.0	0.7
PceGA59	Cherry	183–195	5	0.7	0.9	0.7	UDAp-408	Apricot	168–172	4	0.5	0.6	0.5
PMS67	Cherry	151–181	8	0.7	0.7	0.6	UDAp-410	Apricot	137–169	12	0.9	0.8	0.8
pchgms1	Peach	179–195	7	0.7	0.4	0.7	UDAp-411	Apricot	103–123	10	0.8	0.6	0.8
pchgms2	Peach	162–190	5	0.5	0.2	0.4	UDAp-413	Apricot	135–196	12	0.8	0.5	0.8
pchgms3	Peach	207–217	7	0.6	0.5	0.6	UDAp-414	Apricot	168–188	8	0.8	1.0	0.8
pchgms4	Peach	168–196	6	0.8	0.9	0.7	UDAp-415	Apricot	164–186	10	0.8	0.9	0.8
pchgms5	Peach	178–192	5	0.5	0.3	0.5	UDAp-416	Apricot	105–125	10	0.8	0.8	0.8
pchcms1	Peach	146–166	6	0.8	0.3	0.8	UDAp-418	Apricot	153–185	9	0.8	0.8	0.8
pchcms2	Peach	187–189	2	0.5	0.7	0.4	UDAp-419	Apricot	158–195	18	0.9	0.9	0.9
pchcms4	Peach	241–287	5	0.8	0.8	0.7	UDAp-420	Apricot	171–197	12	0.9	1.0	0.8
pchcms5	Peach	241–301	12	0.8	0.9	0.8	UDAp-421	Apricot	128–174	10	0.7	1.0	0.7
									Average	7	0.7	0.6	0.7

n: number of alleles, He: expected heterozygosity, Ho: observed heterozygosity, PIC: polymorphism information content. *: Tetraploid cherry**.** Source articles for ssrPaCITA, PM and Pce, pchgms, UDA, aprigms and UDAp series are Lopes et al. (2002), Cantini et al. (2001), Sosinski et al. (2000), Testolin et al. (2004), Vilanova et al. (2006), and Messina et al. (2004), respectively.

the least (6, 0.059%) of the total number of dinucleotides. The dinucleotide AG/CT motifs are known to be the most common motifs in plants, and the CG motif is very rare in plants (Varshney et al., 2005). Our findings are consistent with the previous report of Dong et al. (2014), who found that the dinucleotides GA/TC (48.17%) are the most common motif in Siberian apricot. Considering trinucleotides, AGG/CTT was the most abundant repeat motif, which is compatible with previous reports that AAG/CTT was the most common motif in *A. duranensis*, *A. batizocoi*, and *A. hypogaea* (Huang et al., 2016), as well in *Picea abies* and *Pinus taeda* (Ranade et al., 2015).

SSR developments in *Prunus* have continued since 2000 (Yamamoto et al., 2002; Xu et al., 2004; Xie et al., 2006; Vendramin et al., 2007; Chen et al., 2014; Dettor et al., 2015; Sorkheh et al., 2016). The novel 72 EST-SSR loci will be useful in apricot breeding and especially in mapping and anchoring parental maps. The new 72 EST-SSRs identified in the present study enlarge the number of EST-SSRs identified in apricot species, including the 34

identified in apricot (Decroocq et al., 2003; Hagen et al., 2004) and 17 identified in Japanese apricot (Li et al., 2010; Wang et al., 2014). Moreover, the new EST-SSRs contain tetra- and trinucleotide repeats presenting high potential for molecular characterization studies, since these kinds of SSR loci produce clear distinguishable peaks during allele sizing.

Comparison of the novel 72 polymorphic EST-SSRs to 54 polymorphic gSSRs shows that the average number of alleles and the PIC value produced by gSSRs (7.47, 0.65) are higher than those produced by EST-SSR (4.07, 0.43). This higher polymorphism level for gSSRs was expected, because it is known that genomic DNA has higher variability than the more conserved transcribed region (Nicot et al., 2004). Another reason for lower polymorphism levels of the new SSRs is more likely to be their repeats motifs. The new EST-SSRs are mostly tetra- and trinucleotide repeats, which were reported to be less polymorphic in other species, too, such as in pistachio (Topçu et al., 2016) and walnut (Ikhsan et al., 2016). Of the total of 88 published for *Prunus*, only 54 (62%) produced polymorphic bands in apricot, showing that not all publicly available primers are useable for breeding programs. Particularly for mapping purposes, there is a scarcity of reliable SSRs since most SSR loci do not show polymorphism for parents. For example, Ruiz et al. (2011) evaluated 187 SSRs and were able to map only 50 of them. Therefore, the number of primers is not enough since not all *Prunus* SSRs are always transferable in *Prunus* species, as our study showed, and more importantly are not polymorphic for parental accessions. Taken together, this shows that more reliable primers should be developed for apricot mapping studies.

In conclusion, a de novo transcriptome assembly for 7 apricot accessions, ('Stark Early Orange' (SEO), 'Hacıhaliloğlu' (HH), 'Perfection', 'Iğdır', 'Roxana', 'Esen1', and 'Esen2') was performed. A high number of SSR loci in the transcriptomes were identified. Seventy-two polymorphic and reliable EST-SSR loci primers were developed and characterized, amplifying 24 diverse apricot accessions. The new EST-SSRs will be useful for molecular breeding researchers in apricot as well as in other *Prunus* fruits. Furthermore, the low transferability rate (62%) of published *Prunus* gSSRs to apricots demonstrated the need for more reliable SSR primers in molecular breeding studies of apricots, especially for reliable molecular map constructions.

Acknowledgment
This research was supported by the General Directorate of Agricultural Research and Policies (GDAR) of the Turkish Ministry of Food, Agriculture, and Livestock (Project Number TAGEM/12/AR-GE/10) and the Research Fund of Erciyes University (Project Number FCD-2015-6069).

References

Aranzana MJ, Garcia-Mas J, Carbó J, Arús P (2002). Development and variability analysis of microsatellite markers in peach. Plant Breed 121: 87-92.

Bai Y, Hu W, Wang M, He J, Tao Y, Huang W, Feng Z (2016). Transcriptomic analysis of developing embryos of apricot (*Prunus armeniaca* L.). Hortic Environ Biotechnol 57: 197-206.

Cantini C, Iezzoni AF, Lamboy WF, Boritzki M, Struss D (2001). DNA fingerprinting of tetraploid cherry germplasm using simple sequence repeats. J Amer Soc Hort Sci 126: 205-209.

Chen C, Bock C, Okie W, Gmitter F, Jung S, Main D, Beckman T, Wood B (2014). Genome-wide characterization and selection of expressed sequence tag simple sequence repeat primers for optimized marker distribution and reliability in peach. Tree Genet Genomes 10: 1271-1279.

Clarke JB, Tobutt KR (2003). Development and characterization of polymorphic microsatellites from *Prunus avium* 'Napoleon'. Mol Ecol Notes 3: 578-580.

Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21: 3674-3676.

Decroocq S, Cornille A, Tricon D, Babayeva S, Chague S, Eyquard JP, Karychev R (2016). New insights into the history of domesticated and wild apricots and its contribution to Plum pox virus resistance. Mol Ecol 19: 4712-4729.

Decroocq SA, Chague P, Lambert G, Roch JM, Audergon F, Geuna R, Chiozzotto D, Bassi L, Dondini S, Tartarini J et al. (2014). Selecting with markers linked to the *PPVres* major QTL is not sufficient to predict resistance to *Plum Pox Virus* (PPV) in apricot. Tree Genet Genomes 10: 1161-1170.

Decroocq V, Favé MG, Hagen L, Bordenave L, Decroocq S (2003). Development and transferability of apricot and grape EST microsatellite markers across taxa. Theor Appl Genet 106: 912-922.

Dettori MT, Micali S, Giovinazzi J, Scalabrin S, Verde I, Cipriani G (2015). Mining microsatellites in the peach genome: development of new long-core SSR markers for genetic analyses in five *Prunus* species. Springerplus 4: 337.

Dondini L, Lain O, Vendramin V, Rizzo M, Vivoli D, Adami M, Guidarelli M, Gaiotti F, Palmisano F, Bazzoni A et al. (2011). Identification of QTL for resistance to Plum pox virus strain M and D in Lito and Harcot apricot cultivars. Mol Breed 79: 289-299.

Dong S, Liu Y, Niu J, Ning Y, Lin S, Zhang Z (2014). De novo transcriptome analysis of the Siberian apricot (*Prunus sibirica* L.) and search for potential SSR markers by 454 pyrosequencing. Gene 544: 220-227.

Ercisli S (2009). Apricot culture in Turkey. Sci Res Essays 4: 715-719.

Gürcan K, Öcal N, Yılmaz KU, Ullah S, Erdoğan A, Zengin Y (2015). Evaluation of Turkish apricot germplasm using SSR markers: genetic diversity assessment and search for *Plum pox virus* resistance alleles. Sci Hortic 193: 155-164.

Gürcan K, Teber S, Ercisli S, Yilmaz KU (2016). Genotyping by sequencing (GBS) in apricots and genetic diversity assessment with GBS-derived single-nucleotide polymorphisms (SNPs). Biochem Genet 54: 854-885.

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M et al. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc 8: 1494-1512.

Habu T, Tao R (2014). Transcriptome analysis of self-and cross-pollinated pistils of Japanese apricot (*Prunus mume* Sieb. et Zucc.). J Jpn Soc Hortic Sci 83: 95-107.

Hagen S, Chaib J, Fady B, Decroocq V, Bouchet P, Lambert P, Audergon JM (2004). Genomic and cDNA microsatellites from apricot (*Prunus armeniaca* L.). Mol Ecol Notes 4: 742-745.

Halasz J, Pedryc A, Ercisli S, Yilmaz KU, Hegedus A (2010). S-genotyping supports the genetic relationships between Turkish and Hungarian apricot germplasm. J Am Soc Hortic Sci 135:410-417.

Hegedus A, Engel R, Abrankó L, Balogh E, Blázovics A, Hermán R, Halasz J, Ercisli S, Pedryc A, Stefanovits-Bányai É (2010). Antioxidant and antiradical capacities in apricot (*Prunus armeniaca* l.) fruits: variations from genotypes, years, and analytical methods. J Food Sci 75: C722-730.

Howad W, Yamamoto T, Dirlewanger E, Testolin R, Cosson P, Cipriani G, Monforte AJ, Georgi L, Abbott AG, Arus P (2005). Mapping with a few plants, using selective mapping for microsatellite saturation of the *Prunus* reference map. Genetics 171: 1305-1309.

Huang L, Wu B, Zhao J, Li H, Chen W, Zheng Y, Ren X, Chen Y, Zhou X, Lei Y et al. (2016). Characterization and transferable utility of microsatellite markers in the wild and cultivated Arachis species. PLoS One 5: 15.

Ikhsan AD, Topçu H, Sütyemez M, Kafkas S (2016). Novel 307 polymorphic SSR markers from BAC-end sequences in walnut (*Juglans regia L.*): effects of motif types and repeat lengths on polymorphism and genetic diversity. Sci Hortic 213: 1-4.

İpek A, Türkmen Ö, Fidan S, İpek M, Karci H (2016). Genetic variation within the purple carrot population grown in Ereğli District in Turkey. Turk J Agric For 40: 570-576.

Lalli DA, Abbott AG, Zhebentyayeva TN, Badenes ML, Damsteegt V, Polák J, Krška B, Salava J (2008). A genetic linkage map for an apricot (*Prunus armeniaca L.*) BC1 population mapping plum pox virus resistance. Tree Genet Genomes 4: 481-493.

Lambert P, Dicenta F, Rubio M, Audergon JM (2007). QTL analysis of resistance to Sharka disease in the apricot (*Prunus armeniaca L.*) 'Polonais', 'Stark Early Orange' F1 progeny. Tree Genet Genomes 3: 299-309.

Li X, Shangguan L, Song C, Wang C, Gao Z, Yu H, Fang J (2010). Analysis of expressed sequence tags from *Prunus mume* flower and fruit and development of simple sequence repeat markers. BMC Genet 13: 66.

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

Lopes MS, Sefc KM, Laimer M, Da Câmara Machado A (2002). Identification of microsatellite loci in apricot. Mol Ecol Notes 2: 24-216.

Martínez-Gómez P, Crisosto C, Bonghi C, Rubio M (2011). New approaches to Prunus transcriptome analysis. Genetica 139: 755-769.

Martínez-Gómez P, Dicenta F, Audergon JM (2000). Behaviour of apricot (*Prunus armeniaca* L.) cultivars in the presence of Sharka (Plum pox potyvirus): a review. Agronomie 20: 407-422.

Messina R, Lain O, Marrazzo MT, Cipriani G, Testolin R (2004). New set of microsatellite loci isolated in apricot. Mol Ecol Notes 4: 432-434.

Nicot N, Chiquet V, Gandon B, Amilhat L, Legeai F, Leroy P, Bernard M, Sourdille P (2004). Study of simple sequence repeat (SSR) markers from wheat expressed sequence tags (ESTs). Theor Appl Genet 109: 800-805.

Niu J, An J, Wang L, Fang C, Ha D, Fu C, Qiu L, Yu H, Zhao H, Hou X (2015). Transcriptomic analysis revealed the mechanism of oil dynamic accumulation during developing Siberian apricot (*Prunus sibirica* L.) seed kernels for the development of woody biodiesel. Biotechnol Biofuels 8: 29.

Ranade SS, Lin YC, De Peer YV, García-Gil MR (2015). Comparative in silico analysis of SSRs in coding regions of high confidence predicted genes in Norway spruce (*Picea abies*) and Loblolly pine (*Pinus taeda*). BMC Genet 16: 149.

Rozen S, Skaletsky H, Krawetz S, Misener S (2000). Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol 132: 365-86.

Rubio M, Ballester AR, Olivares PM, Castro de Moura M, Dicenta F, Martínez-Gómez P (2015). Gene expression analysis of plum pox virus (*Sharka*) susceptibility/resistance in apricot (*Prunus armeniaca* L.). PLoS One 10: e0144670.

Rubio M, Ruiz D, Egea J, Martínez-Gómez P, Dicenta F (2014). Opportunities of marker assisted selection for Plum pox virus resistance in apricot breeding programs. Tree Genet Genomes 10: 513-525.

Ruiz EMV, Soriano JM, Romero C, Zhebentyayeva T, Terol J, Zuriaga E, Llácer G, Abbott AG, Badenes ML (2011). Narrowing down the apricot *Plum pox virus* resistance locus and comparative analysis with the peach genome syntenic region. Mol Plant Pathol 12: 535-547.

Sakar E, Ünver H (2016). Molecular characterization of ancient olive genotypes from Hatay province in Turkey. Turk J Agric For 40: 795-801.

Salazar JA, Rubio M, Ruiz D, Tartarini S, Martínez-Gómez P, Dondini L (2015). SNP development for genetic diversity analysis in apricot. Tree Genet Genome 11: 15.

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

Soriano JM, Domingo ML, Zuriaga E, Romero C, Zhebentyayeva T, Abbott A, Badenes ML (2012). Identification of simple sequence repeat markers tightly linked to Plum pox virus resistance in apricot. Mol Breed 30: 1017-1026.

Sorkheh K, Khaleghi E (2016). Molecular characterization of genetic variability and structure of olive (*Olea europaea* L.) germplasm collection analyzed by agromorphological traits and microsatellite markers. Turk J Agric For 40: 583-596.

Sorkheh K, Prudencio AS, Ghebinejad A, Dehkordi MK, Erogul D, Rubio M, Martínez-Gómez P (2016). In silico search, characterization and validation of new EST-SSR markers in the genus Prunus. BMC Res Notes 9: 336.

Sosinski B, Gannavarapu M, Hager LE, Beck LE, King GJ, Ryder CD, Rajapakse S, Baird WV, Ballard RE, Abbott AG (2000). Characterization of microsatellite markers in peach [*Prunus persica* (L) Batsch]. Theor Appl Genet 101: 421-428.

Testolin R, Messina R, Lain O, Marrazzo MT, Huang WG, Cipriani G (2004). Microsatellites isolated in almond from an AC-repeat enriched library. Mol Ecol Notes 4: 459-461.

Topçu H, Çoban N, Kafkas S (2016). Novel microsatellite markers in *Pistacia vera* L. and their transferability across the genus *Pistacia*. Sci Hortic 198: 91-97.

Van Bel M, Proost S, Van Neste C, Deforce D, Van de Peer Y, Vandepoele K (2013). TRAPID: An efficient online tool for the functional and comparative analysis of de novo RNA-Seq transcriptomes. Genome Biol 14: 134.

Varshney RK, Graner A, Sorrells ME (2005). Genic microsatellite markers in plants: features and applications. Trends Biotechnol 23: 48-55.

Vavilov NI (1951). Phytogeographical basis of plant breeding. The origin, variation, immunity and breeding of cultivated plants. Chronica Botanica 13: 13-54.

Vendramin E, Dettori MT, Giovinazzi J, Micali R, Quarta R, Verde I (2007). A set of EST-SSRs isolated from peach fruit transcriptome and their transportability across *Prunus* species. Mol Ecol Notes 7: 307-310.

Vilanova S, Soriano JM, Lalli DA, Romero C, Abbott AG, Llácer G, Badenes ML (2006). Development of SSR markers located in the G1 linkage group of apricot (*Prunus armeniaca*) using a bacterial artificial chromosome library. Mol Ecol Notes 6: 789-791.

Wang YJ, Li XY, Han J, Fang WM, Li XD, Wang SS (2014). Analysis of genetic relationships and identification of flowering-mei cultivars using EST-SSR markers developed from apricot and fruiting-mei. Sci Hortic 132: 12-17.

Wang X, Lu P, Luo Z (2013). GMATo: a novel tool for the identification and analysis of microsatellites in large genomes. Bioinformation 9: 541-544.

Xie H, Sui Y, Chang FQ, Xu Y, Ma RC (2006). SSR allelic variation in almond (*Prunus dulcis* Mill.). Theor Appl Genet 112: 366-372.

Xu Y, Ma RC, Xie H, Liu JT, Cao MQ (2004). Development of SSR markers for the phylogenetic analysis of almond trees from China and the Mediterranean region. Genome 47: 1091-1104.

Yamamoto T, Mochida K, Imai T, Shi YZ, Ogiwara I, Hayashi T (2002). Microsatellite markers in peach [*Prunus persica* (L.) Batsch] derived from an enriched genomic and cDNA libraries. Mol Ecol Notes 2: 298-301.

Zhong W, Gao Z, Zhuang W, Shi T, Zhang Z, Ni Z (2013). Genome-wide expression profiles of seasonal bud dormancy at four critical stages in Japanese apricot. Plant Mol Biol 83: 247-264.