Characterization of hawthorn (*Crataegus* spp.) genotypes by SSR markers

Murat Güney¹ · Salih Kafkas² · Hakan Keles¹ · Servet Aras¹ · Sezai Ercişli³

Received: 13 May 2018 / Revised: 17 September 2018 / Accepted: 17 September 2018 / Published online: 28 September 2018
© Prof. H.S. Srivastava Foundation for Science and Society 2018

Abstract Hawthorn (*Crataegus* spp.) is an edible wild fruit that is used in traditional medicine, landscape studies, and food and beverage industries in many countries. It is an important wild plant species in Turkey and is numerous in the Yozgat Province. Genetic and breeding studies on hawthorn are very limited. Therefore, we aimed to characterize 91 hawthorn genotypes using simple sequence repeat (SSR) markers. The SSRs were developed from apple and pear and were screened in hawthorn for amplification and polymorphisms. A total of 265 alleles were detected from thirty-two SSR primer pairs, and those were used to identify genetic relationships. The number of alleles ranged from 2 to 21 alleles per locus with a mean value of 8.28. The Hi05b09 locus showed the highest allele number ($N_a = 21$). The polymorphism information content (PIC) values ranged from 0.16 (CH03d10) to 0.89 (C6554) with a mean value of 0.60. An Unweighted Pair Group Method with Arithmetic Average method was used to cluster the genotypes, and four major clusters were obtained from the amplification of the SSRs. STRUCTURE software identified four populations ($\Delta K = 4$) and eight sub-populations ($\Delta K = 8$), and four major clusters similar results to UPGMA analysis. Our study showed that the SSR markers could be utilized as a reliable tool for the determination of genetic variations and relationships of

hawthorn genotypes. A basic molecular analysis on the hawthorn genotypes identified in this study will promote the collection of germplasm collection and the selection of parents' in future cross-breeding studies.

Keywords Hawthorn · *Crataegus* spp. · SSR · Structure

Introduction

Plants are a source of natural wealth to countries worldwide in the form of oil and other underground sources (Nas 2012). There are 1225 species and 9000 taxa that belong to 163 families in Turkey's flora, of which 3000 are endemic species. This indicates that Turkey is wealthy in plant resources (Özgen et al. 2000). Hawthorn is a wild species that grows naturally in the Aegean and the eastern, southern, central, and northern Anatolia regions (Dönmez 2004; Ercişli 2004).

The genus *Crataegus* ranges from small shrubs to trees that are yellow to black, it produces an important wild fruit, and there are 200 species worldwide with 21 species occurring in Turkey. Hawthorn is a common edible wild fruit in Turkey that has rich nutrients important for human health and nutrition. It is a genetically important pome fruit that grows in the different ecological conditions of Turkey. Hawthorn is not only cultivated in Turkey but is also used traditionally by herbalists, in landscape studies, and in the food and beverage industries in other countries.

Hawthorn breeding studies are limited in Turkey. In the literature, some characteristics of hawthorn genotypes, such as fruit size, plant habitat, leaf properties, and phytochemical substances, have been determined, however, these studies are limited, and many genetic materials could not be controlled. Loss of genetic material is easy, but its

✉ Murat Güney
murat.guney@bozok.edu.tr

¹ Department of Horticulture, Faculty of Agriculture, University of Yozgat Bozok, Yozgat, Turkey

² Department of Horticulture, Faculty of Agriculture, University of Cukurova, Adana, Turkey

³ Department of Horticulture, Faculty of Agriculture, University of Atatürk, Erzurum, Turkey

recovery is almost impossible. The selection and preservation of genetic material is of importance for future breeding and development studies. The selection, propagation, and preservation of genotypes are possible with many methods, including current molecular breeding methods. Simple sequence repeats (SSRs) are widely used because they are codominant, suitable for automation, highly polymorphic, rapid, and reliable. The previous literature is limited on data regarding the selection and characterization of hawthorn species.

Molecular markers are used to determine the genetic relationships within plant populations with almost 100% reliability. Additionally, molecular markers are utilized in plant systematics, plant breeding, and the evaluation of genetic resources (Gülşen and Mutlu 2005). Many molecular methods are used in characterization studies, such as Restriction Fragment Length Polymorphism (RFLP), Random Amplified Polymorphic DNA (RAPD), Inter-Simple Sequence Repeats (ISSR), Amplified Fragment Length Polymorphism (AFLP), and SSRs (Tanksley et al. 1989; Williams et al. 1990; Zietkiewicz et al. 1994; Vos et al. 1995; Kafkas et al. 2006a). Although RAPD and ISSR techniques were commonly used in previous molecular characterization studies for the genus *Crataegus*, SSR methods are more useful than RAPD and ISSR for studying plant genetics and breeding. SSR markers are highly polymorphic, codominant, dependable, and convenient. In many previous studies, SSR markers were used for the molecular characterizations of many plants, such as apricot (Wang et al. 2011), strawberry (Govan et al. 2008), apple (Potts et al. 2012), pistachio (Zaloglu et al. 2015), and walnut (Ikhsan et al. 2016).

A study on this less-common cultivar is of interest because of its potential importance in horticulture. The objective of this study was to complete the molecular characterization of individual hawthorn genotypes that are maintained in the Yozgat Province of Turkey using SSR markers, which will help establish the potential sources of breeding material for future uses of hawthorn.

Materials and methods

Plant material

Exactly, 91 different genotypes of *Crataegus* spp. collected from the Yozgat province of Turkey were used as study material (Table 1).

DNA extraction

Extraction of total genomic DNA of frozen leaves was done by the CTAB method following the protocol

described by Doyle and Doyle (1987) with minor modifications (Kafkas et al. 2006b). The DNA yield and quality were measured by Qubit Fluorometer (Invitrogen) using standard operating protocols. The isolated DNA was subsequently diluted to 10 ng/μl for SSR-PCR reactions and stored at − 20 °C.

SSR-PCR reactions

One-hundred SSR primers derived from apples and pears were used to screen eight genotypes of hawthorn. A set of 32 polymorphic SSRs was selected for the characterization of 91 hawthorn genotypes (Table 2).

All SSR-PCR reactions were done based on a three-primer strategy according to Scheulke (2000) with minor modifications. The reactions were done in a total volume of 12.5 μl containing 10 ng DNA, 75 mM Tris-HCl (pH 8.8), 20 mM (NH₄)₂SO₄, 2.0 mM MgCl₂, 0.01% Tween 20, 200 μM of each dNTP, 10 nM M13 tailed forward primer at the 5' end, 200 nM reverse primer, 200 nM universal M13 tail primer (5'-TGTAACGACGCGCCAGT-3') labeled with FAM, VIC, NED or PET dye, and 0.6 U of Taq DNA polymerase. PCR amplifications were carried out in two successive steps. The first step involved initial denaturation at 94 °C for 3 min, followed by 28 cycles of 94 °C for 30 s, 58 °C for 45 s and 72 °C for 60 s. The second step included 10 cycles of 94 °C for 30 s, 52 °C for 45 s and 72 °C for 60 s, and a final extension at 72 °C for 5 min. The completed PCR reactions were denatured for use in capillary electrophoresis in an ABI 3130xl genetic analyzer [Applied Biosystems Inc., Foster City, CA, USA (ABI)] using a 36-cm capillary array with POP7 as the matrix (ABI). In order to cause denaturation of the samples, 0.5 μl (in 6-FAM and VIC labeled primers) or 1.0 μl (in NED and PET labeled primers) of the amplified product were mixed with 0.3 μl of the size standard and 9.7 μl of Hi-Di formamide. The fragments were resolved using the ABI data collection software v3.0, and SSR fragment analysis was performed with the Genmap v4.0 (Applied Biosystems Inc.).

Data analysis

After capillary electrophoresis of the SSR loci, the number of alleles per locus (*Na*), effective number of alleles (*Ne*), expected heterozygosity (*He*) and observed heterozygosity (*Ho*) were computed using the GenAlEx version 6.5 program (Peakall and Smouse 2012). Calculation of polymorphism information content (*PIC*) for the loci was done by PowerMarker software version 3.25 (Liu and Muse 2005). A dendrogram was obtained using Jaccard coefficient in NTSYSpc v2.21c (Rohlf 2009) software by an

Table 1 The origin of hawthorn genotypes used in this study

No	Genotype	Location	No	Genotype	Location
1	Ağılı1	Aydıncık/Yozgat	47	Çekerek1	Çekerek/Yozgat
2	Ağılı2	Aydıncık/Yozgat	48	Çekerek2	Çekerek/Yozgat
3	Ağılı3	Aydıncık/Yozgat	49	Çekerek3	Çekerek/Yozgat
4	Ağılı4	Aydıncık/Yozgat	50	Çekerek4	Çekerek/Yozgat
5	Akbucak1	Sarıkaya/Yozgat	51	Çekerek6	Çekerek/Yozgat
6	Akbucak2	Sarıkaya/Yozgat	52	Eymir1	Sorgun/Yozgat
7	Akdağ1	Akdağmadeni/Yozgat	53	Eymir2	Sorgun/Yozgat
8	Akdağ2	Akdağmadeni/Yozgat	54	Eymir3	Sorgun/Yozgat
9	Akdağ3	Akdağmadeni/Yozgat	55	Eymir4	Sorgun/Yozgat
10	Akdağ5	Akdağmadeni/Yozgat	56	Eymir5	Sorgun/Yozgat
11	Akdağ6	Akdağmadeni/Yozgat	57	Eymir6	Sorgun/Yozgat
12	Akdağ7	Akdağmadeni/Yozgat	58	Hattuşaş 1	Merkez/Yozgat
13	Akdağ8	Akdağmadeni/Yozgat	59	Hattuşaş 2	Merkez/Yozgat
14	Albatros1	Merkez/Yozgat	60	Kadısehri7	Kadısehri/Yozgat
15	Albatros2	Merkez/Yozgat	61	Kadısehri1	Kadısehri/Yozgat
16	Albatros3	Merkez/Yozgat	62	Kadısehri2	Kadısehri/Yozgat
17	Albatros4	Merkez/Yozgat	63	Kadısehri3	Kadısehri/Yozgat
18	Albatros6	Merkez/Yozgat	64	Kadısehri4	Kadısehri/Yozgat
19	Albatros7	Merkez/Yozgat	65	Kadısehri5	Kadısehri/Yozgat
20	Albatros8	Merkez/Yozgat	66	Kadısehri6	Kadısehri/Yozgat
21	Alemdar4	Merkez/Yozgat	67	Karaelli1	Sarıkaya/Yozgat
22	Alemdar6	Merkez/Yozgat	68	Karaelli2	Sarıkaya/Yozgat
23	Azizli1	Merkez/Yozgat	69	Karaelli3	Sarıkaya/Yozgat
24	Azizli2	Merkez/Yozgat	70	Karaelli4	Sarıkaya/Yozgat
25	Baraj1	Merkez/Yozgat	71	Kürkçü1	Sarıkaya/Yozgat
26	Baraj2	Merkez/Yozgat	72	Kürkçü2	Sarıkaya/Yozgat
27	Baraj3	Merkez/Yozgat	73	Kürkçü4	Sarıkaya/Yozgat
28	Baraj4	Merkez/Yozgat	74	Kürkçü5	Sarıkaya/Yozgat
29	Baraj6	Merkez/Yozgat	75	Kürkçü6	Sarıkaya/Yozgat
30	Baraj7	Merkez/Yozgat	76	Lök 4	Merkez/Yozgat
31	Battal	Merkez/Yozgat	77	Lök1	Merkez/Yozgat
32	Büyükören	Sorgun/Yozgat	78	Lök2	Merkez/Yozgat
33	Cezaevi1	Merkez/Yozgat	79	Lök3	Merkez/Yozgat
34	Cezaevi2	Merkez/Yozgat	80	Lök5	Merkez/Yozgat
35	Cezaevi3	Merkez/Yozgat	81	Poyrazlı	Boğazlıyan/Yozgat
36	Cezaevi4	Merkez/Yozgat	82	Sırçalı	Boğazlıyan/Yozgat
37	Cezaevi5	Merkez/Yozgat	83	Sırçalı2	Boğazlıyan/Yozgat
38	Çamlık1	Merkez/Yozgat	84	Şahintepe1	Merkez/Yozgat
39	Çamlık2	Merkez/Yozgat	85	Şahintepe2	Merkez/Yozgat
40	Çamlık3	Merkez/Yozgat	86	Şahintepe3	Merkez/Yozgat
41	Çamlık4	Merkez/Yozgat	87	Şahintepe4	Merkez/Yozgat
42	Çamlık5	Merkez/Yozgat	88	Şahintepe5	Merkez/Yozgat
43	Çamlık6	Merkez/Yozgat	89	Şahintepe6	Merkez/Yozgat
44	Çandır1	Çandır/Yozgat	90	Topaktaş	Sorgun/Yozgat
45	Çayıralan1	Çayıralan/Yozgat	91	Yenipazar	Boğazlıyan/Yozgat
46	Çayıralan3	Çayıralan/Yozgat			

unweighted pair-group method with arithmetic averages (UPGMA).

Population structure and identification of admixed individuals was performed using the model-based software

Table 2 Summary of SSR primers developed from apple and pear tested in hawthorn genotypes

No	Origin and reference	Acronyms	No. of tested loci	No. of non-amplified loci	Monomorphic	Polymorphic	Transferability rate (%)	Polymorphic rate (%)
1	Guilford et al. (1997)	Nz	1	–	–	1	100	100
2	Hokanson et al. (1998)	GD	3	–	2	1	100	33.3
3	Testolin et al. (2000)	UDP ^a	1	–	–	1	100	100
4	Oddou-Muratorio et al. (2001)	MSS	2	–	–	2	100	100
5	Liebhart et al. (2002)	CH	29	13	5	11	55.1	37.9
6	Yamamoto et al. (2002)	BG, HG, KA, KB, KU	5	–	5	–	100	–
7	Hemmat et al. (2003)	GD	5	2	1	2	60	40
8	Silfverberg-Dilworth et al. (2006)	CN, Hi	28	8	13	7	71.4	25
9	Fernandez-Fernandez et al. (2006)	EMPC ^a	4	1	1	2	75	50
10	Inoue et al. (2007)	IPPN	1	1	–	–	–	–
11	Han et al. (2009)	BAC	4	3	1	–	25	–
12	Han et al. (2011)	CN, CTG	13	7	3	3	46.1	23.1
13	Moriya et al. (2012)	MEST	1	–	–	1	100	100
14	Wang et al. (2012)	C	3	1	1	1	66.7	33.3
Total			100	36	32	32	64	32

^aSSRs isolated from pear

program, STRUCTURE v2.3.4 (Pritchard et al. 2000) in which a number of populations (K) are assumed to be present that is characterized by a set of allele frequencies at each locus. Individuals in the sample are assigned to populations (clusters), or jointly to more populations if their genotypes indicate that they are admixed. Delta K determines probable population number using Ln P (D) values (logarithm probability for each K). The term Delta K is calculated by change ratio of logarithm probability ($\Delta K = 2$ to $\Delta K = 10$). The information about probable population number was shown by the highest K of Delta K in the diagram.

Results

In the current study, 32 of 100 SSR polymorphic primer pairs generated scorable and polymorphic bands that were subsequently used for fingerprinting all of the hawthorn genotypes.

Polymorphism levels of SSR loci

Among the 100 SSR primer pairs utilized to fingerprint the 91 hawthorn genotypes, 32 SSR markers produced a monomorphic band, 36 SSR markers failed to amplify specific DNA targets, and the remaining 32 SSR markers

generated scorable polymorphic bands that were subsequently used for characterization of the hawthorn genotypes (Table 2). The transferability rate of the 100 SSR loci was 64%, while the polymorphism rate was 50% in the amplified loci.

Among the 91 genotypes, a total of 265 alleles were detected. The number of alleles revealed by the SSR analysis ranged from 2 to 21 alleles per locus (Table 3) with a mean value of 8.28 alleles per locus. The highest allelic number ($N_a = 21$) was given by the Hi05b09 locus. The number of effective alleles ranged from 1.21 (CH03d10) to 9.46 (C6554) with a mean of 3.84 (N_e). The observed heterozygosity varied from 0.00 to 1.00 with a mean of 0.65. Observed heterozygosity was determined to be higher in the CTG105511 and the Hi21e04 loci compared to other loci. The average value of expected heterozygosity was 0.64, and the highest value (0.89) was observed in the C6554 locus. The polymorphism information content values ranged from 0.16 (CH03d10) to 0.89 (C6554) with a mean PIC value of 0.60 (Table 3).

Genetic relationships among hawthorn genotypes

The UPGMA algorithms that the dendrogram generated, revealed the relationships among the genotypes and are shown in Fig. 1. The genetic similarity coefficients ranged from 0.38 to 1.00. The dendrogram was grouped into four

Table 3 Allele range, number of alleles (*Na*), number of effective alleles (*Ne*), observed heterozygosity (*Ho*), expected heterozygosity (*He*) and polymorphism information content (*PIC*) values based on 32 SSR loci developed from *Crataegus* spp.

No.	SSR loci	Allele range	Na	Ne	Ho	He	PIC
1	C6554	283–328	19	9.46	0.79	0.89	0.89
2	CH01f12	230–240	4	2.33	0.5	0.57	0.48
3	CH02a04	102–126	8	4.16	0.83	0.76	0.72
4	CH02b10	129–151	9	6.62	0.73	0.85	0.83
5	CH02d11	118–127	4	2.39	0.38	0.58	0.5
6	CH02f06	185–228	13	7.11	0.97	0.86	0.84
7	CH03d10	180–182	2	1.21	0.19	0.17	0.16
8	CH03e03	188–225	8	4.59	0.9	0.78	0.75
9	CH03g12	173–211	14	6.63	0.99	0.85	0.83
10	CH04d08	129–161	15	7.74	0.81	0.87	0.86
11	CH05c02	168–195	8	4.98	0.78	0.8	0.77
12	CH05g07	169–195	11	5.26	0.77	0.81	0.79
13	CTG1057511	144–180	8	3.6	1	0.72	0.69
14	CTG1064471	204–233	8	3.51	0.92	0.72	0.68
15	CTG1074058	233–244	4	1.73	0.59	0.42	0.34
16	EMPC102 ^a	55–159	3	1.62	0.46	0.38	0.35
17	EMPC116 ^a	170–182	4	1.99	0.54	0.5	0.43
18	GD-100	243–245	2	1.98	0	0.5	0.37
19	GD-136	101–108	4	1.9	0.49	0.47	0.44
20	GD-154	180–216	14	7.02	0.94	0.86	0.84
21	Hi03a03	179–206	7	5.3	0.98	0.81	0.78
22	Hi04a08	194–198	2	1.73	0.56	0.42	0.33
23	Hi04c10	174–203	8	2.4	0.22	0.58	0.52
24	Hi05b09	150–245	21	5.87	0.56	0.83	0.81
25	Hi08h12	118–156	12	5.05	0.99	0.8	0.78
26	Hi11a01	194–203	4	1.61	0.41	0.38	0.35
27	Hi21e04	170–205	18	4.31	1	0.77	0.74
28	MEST187	214–237	9	2.51	0.64	0.6	0.57
29	MSSR3	159–165	2	1.61	0.51	0.38	0.31
30	MSSR4	286–299	4	1.64	0.39	0.39	0.33
31	Nz17e6	98–114	7	3.65	0.72	0.73	0.69
32	UDP98-416 ^a	107–136	9	1.46	0.31	0.31	0.31
	Total	–	265	–	–	–	–
	Min	–	2	1.21	0	0.17	0.16
	Mean	–	8.28	3.84	0.65	0.64	0.6
	Max	–	21	9.46	1	0.89	0.89

^aSSRs isolated from pear

major clusters. All of the clusters included the genotypes among 1–30 (Cluster-I), 31–45 (Cluster-II), 46–83 (Cluster-III), and 84–91 (Cluster-IV) (Fig. 1).

Thirty genotypes belonged to Cluster-I, and the Karaelli1 and Kadişehri2, Karaelli4 and Azizli1, and Şahintepe1 and Şahintepe2 pairs were identical (Fig. 1). The lowest genetic similarity coefficient (0.43) was scored

by the Kürkcü1 and Albatros7 genotype pair in Cluster-I. Moreover, Şahintepe3 and Albatros1 were found to be very close to each other with a 0.94 similarity value.

Fifteen genotypes were in Cluster-II, and the Cezaevi1 and Cezaevi4, and Alemdar6 and Çamlık5 pairs displayed the highest genetic similarity coefficient (0.95). The lowest (0.33) value was scored by the Azizli2 and Büyükören genotype pair. Furthermore, Hattuşaş2 and Azizli2 were closely related with a 0.91 similarity value.

Thirty-eight genotypes were in Cluster-III. Several pairs of genotypes could not be separated within Cluster-III. The highest genetic similarity coefficient (0.85) was obtained by the Ağılı2 and Çekerek1 genotype pair. The lowest (0.35) similarity value was scored by the Baraj7 and Kadişehri4 genotype pair in Cluster-III.

Eight genotypes were in Cluster-IV. The highest genetic similarity coefficient (0.96) was scored by the Poyrazlı and Sırçalı2 genotype pair. The lowest (0.59) similarity value was scored by the genotype pairs of Lök1 and Topaktaş and Lök1 and Akbucak1.

Structure analysis

Similar to the NTSYSpc v2.21c, the structural genetic analysis was performed on the 91 hawthorn genotypes using 32 amplified loci with the STRUCTURE and Structure Harvester programs. All the genotypes were divided into four main clusters, similar to the results of the UPGMA analysis (Fig. 1). The highest Delta K (ΔK) value occurred at $\Delta K = 4$ (Fig. 2), which corresponded to the most probable number of populations, and $\Delta K = 8$ corresponded to the number of sub-populations in the study (Figs. 2, 3). According to $\Delta K = 4$, the 1–30 genotypes grouped in Cluster-I, 31–45 were in Cluster-II, 46–83 were in Cluster-III, and 84–91 were in Cluster-IV. Furthermore, in $\Delta K = 4$, genotypes 34, 41, 79, and 91 had mutual alleles inside and outside of the assigned clusters (Figs. 1, 3).

Discussion

SSR markers are a fast and reliable DNA-based marker technique. Thirty-two polymorphic SSR primer pairs, including 3 SSRs from pear and 29 SSRs from apple, were used in this study. Zhang et al. (2008), Khiari et al. (2015) and Emami et al. (2018) also used SSRs developed from apple and pear. Few studies in the literature have evaluated the molecular characterization of hawthorn genotypes (Zhang et al. 2008; Dai et al. 2009; Rajeb et al. 2010; Yilmaz et al. 2010; Beigmohamadi and Rahmani 2011; MirAli et al. 2011; Serce et al. 2011; Khiari et al. 2015; Rahmani et al. 2015; Erfani-Moghadam et al. 2016; Fang et al. 2017; Emami et al. 2018).

Fig. 1 UPGMA and STRUCTURE ($\Delta K = 4$) analysis of *Crataegus* spp. genotypes

Fig. 2 Value of ΔK estimated for the structure analysis of *Crataegus* spp. genotypes ($\Delta K = 4$ populations and $\Delta K = 8$ sub-populations)

In the current study, we present polymorphic SSR loci for hawthorn that were developed from apple and pear (Table 3). The transferability of the tested SSR loci was

relatively high. Gasic et al. (2009) reported the transferability of apple SSRs across Rosaceae ranged from 25% in apricot to 59% in pear. The transferability of pear SSRs across Rosaceae was also studied by Fan et al. (2013) and ranged from 58.2% in apple to 11.9% in cherry. In addition, Lo et al. (2009) identified 23 SSRs from apple, and 13 (56.5%) were transferable to hawthorn. The transferability ratio was higher in our study than the previously mentioned studies. Additionally, the rate of polymorphisms within the transferable hawthorn SSRs was lower in other studies than the 32% calculated in this study.

Many studies report that SSR loci with highly relevant transferability can be useful in estimating the molecular fingerprinting within the Rosaceae family (i.e. *Crataegus* spp.). In a study by Zhang et al. (2008), 100 polymorphic SSR loci were screened for 37 hawthorn genotypes, and a total of 91 alleles were detected. The number of alleles revealed by the SSR analysis ranged from 3 to 13 alleles per locus with a mean value of 9.1 alleles per locus. Lo et al. (2009) reported 13 SSRs for 239 individuals and obtained a total of 251 alleles (7–34 per locus) with a mean value of 19.3. Beigmohamadi and Rahmani (2011) used 9

Fig. 3 Population structure of 91 *Crataegus* spp. genotypes estimated from 32 SSRs using STRUCTURE program ($\Delta K = 2$ to $\Delta K = 10$)

RAPD primers for 5 hawthorn genotypes, and 101 alleles were revealed with a mean value of 11.2 (N_a). Rahmani et al. (2015) reported 18 ISSR loci for 164 hawthorn genotypes and found 254 alleles with a mean value of 14.11. In an experiment by Khiari et al. (2015), 7 SSR loci were tested for 86 hawthorn genotypes, and the mean value of alleles was 2.89. Furthermore, Fang et al. (2017) studied 12 ISSR loci for 92 hawthorn genotypes and identified 298 alleles with a mean value of 25.25. Emami et al. (2018) published a total of 76 alleles from 11 SSR primers in the characterization of 201 hawthorn accessions. In our study, 265 alleles with a mean value of 8.28 were identified by 32 SSR primers in 91 hawthorn genotypes.

The PIC was calculated as 0.60 in the current study but Khiari et al. (2015) and Emami et al. (2018) reported 0.89 and 0.60, respectively. In this study, H_e and H_o were 0.64 and 0.65, respectively, but Khiari et al. (2015) reported 0.66 and 0.42, and Emami et al. (2018) determined H_e and H_o to be 0.64 and 0.83, respectively. The PIC and H_e values in current study and those in Emami et al. (2018) were lower than the values obtained by Khiari et al. (2015). H_o was also found to be lower in this study and Khiari et al. (2015) than the values reported by Emami et al. (2018). These lower values may be the result of using different hawthorn genotypes and SSR primer numbers between the studies.

Moreover, the SSR-based structural genetic analysis and the UPGMA clustering produced similar results on the genetic relationships of the hawthorn genotypes. We used a total of 32 polymorphic SSR pairs, in which only 2 markers were used in earlier studies, namely CH02b10 in Zhang et al. (2008) and CH05g07 in Khiari et al. (2015) and Emami et al. (2018). The remaining 30 markers were used for the first time in the current study. The clusters formed from all 32 markers. The markers with highest allele numbers and PIC values that were obtained from STRUCTURE and UPGMA data analysis can be used to identify genotypes. The highest number of alleles was obtained from the Hi05b09 loci, and the highest PIC value was obtained from C6554. The Hi05b09 and C6554 markers can be used further to investigate hawthorn genotypes. Our analyses prove that several genotypes are identical even though they were gathered from different geographical locations. Based on a genetic investigation, it is possible that Kadışehri2 and Karaelli1 are ΔK the same genotype. Similarly, Karaelli4 and Azizli1 and Şahintepe1 and Şahintepe2 may be the same genotype (Table 1 and Fig. 1). In addition, some genotypes such as Hattuşas2, Kürkcü5, Eymir5, Çamlık1, and Lök1 were close within the clusters and at the out-cluster genotypes. Several genotype pairs could not be separated, and Cezaevi3 was in the out-group of Cluster-III.

Conclusion

Despite the abundance of genetic sources of hawthorn in Turkey, this is the first molecular characterization study using SSR markers for hawthorn. There is very limited information in the literature on the genetics of hawthorn. We used SSR markers that were developed from apple and pear, which had satisfactory transferability and polymorphisms for the characterization of hawthorn genotypes. Novel SSR marker development in a species requires sequence information, and it is necessary to follow an expensive procedure to reveal the polymorphic SSR markers. Therefore, this study presents polymorphic SSRs for future genetic studies on hawthorn. A phenotypic characterization of the genotypes may also initiate future association mapping studies on hawthorn to detect the markers associated with economically important characters. Finally, the results of this study will help to manage hawthorn germplasm collections and assist with the decision on choosing parents for future cultivar breeding programs.

Acknowledgements The authors thank the Bozok University Scientific Research Projects Unit (Project No. 6602c-ZF/17-95) for financial support.

Compliance with ethical standards

Conflict of interest The authors state that they do not have any conflict of interest.

References

- Beigomohadi M, Rahmani F (2011) Genetic variation in hawthorn (*Crataegus* spp.) using RAPD markers. *Afr J Biotechnol* 10(37):7131–7135
- Dai H, Guo X, Zhang Y, Li Y, Chang L, Zhang Z (2009) Application of random amplified polymorphic DNA and inter-simple sequence repeat markers in the genus *Crataegus*. *Ann Appl Biol* 154:175–181. <https://doi.org/10.1111/j.1744-7348.2008.00290.x>
- Dönmez AA (2004) The genus *Crataegus* L. (Rosaceae) with special reference to hybridisation and biodiversity in Turkey. *Turk J Bot* 28:29–37
- Doyle JJ, Doyle JL (1987) A Rapid isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull* 19:11–15
- Emami A, Shabanian N, Rahmani MS, Khadivi A, Mohammad-Panah N (2018) Genetic characterization of the *Crataegus* genus: implications for in situ conservation. *Sci Hortic* 213:56–65. <https://doi.org/10.1016/j.scienta.2017.12.014>
- Ercisli S (2004) A short review of the fruit germplasm resources of Turkey. *Genet Resour Crop Evol* 51:419–435. <https://doi.org/10.1023/B:GRES.0000023458.60138.79>
- Erfani-Moghadam J, Mozafari M, Fazeli A (2016) Genetic variation of some hawthorn species based on phenotypic characteristics and RAPD marker. *Biotechnol Biotechnol Equip* 30(2):247–253. <https://doi.org/10.1080/13102818.2015.1121790>
- Fan L, Zhang M, Liu Q, Li L, Song Y, Wang L, Zhang S, Wu J (2013) Transferability of newly developed pear SSR markers to other

- Rosaceae Species. *Plant Mol Biol Rep* 31(6):1271–1282. <https://doi.org/10.1007/s11105-013-0586-z>
- Fang S, Shu-ying C, Jia T, Peng L, Xue Q, Lei W, Shu-ping L, Jiang L (2017) Morphological and ISSR molecular markers reveal genetic diversity of wild hawthorns (*Crataegus songorica* K. Koch.) in Xinjiang China. *J Integr Agric* 16(11):2482–2495. [https://doi.org/10.1016/S2095-3119\(17\)61688-5](https://doi.org/10.1016/S2095-3119(17)61688-5)
- Fernandez-Fernandez F, Harvey NG, James CM (2006) Isolation and characterization of polymorphic microsatellite markers from European pear (*Pyrus communis* L.). *Mol Ecol Notes* 6:1039–1041. <https://doi.org/10.1111/j.1471-8286.2006.01422.x>
- Gasic K, Han Y, Kertbundit S, Shulaev V, Iezzoni AF, Stover EW, Bell RL, Wisniewski ME, Korban SS (2009) Characteristics and transferability of new apple EST-derived SSRs to other Rosaceae species. *Mol Breed* 23:397–411. <https://doi.org/10.1007/s11032-008-9243-x>
- Govan CL, Simpson DW, Johnson AW, Tobutt KR, Sargent DJ (2008) A reliable multiplexed microsatellite set for genotyping *Fragaria* and its use in a survey of 60 *F. × ananassa* cultivars. *Mol Breed* 22(4):649–661. <https://doi.org/10.1007/s11032-008-9206-2>
- Guilford P, Prakash S, Zhu JM, Rikkerink E, Gardiner S, Bassett H, Forster R (1997) Microsatellites in *Malus × domestica* (apple): abundance, polymorphism and cultivar identification. *Theor Appl Genet* 94:249–254. <https://doi.org/10.1007/s001220050407>
- Gülşen O, Mutlu N (2005) Bitki biliminde kullanılan genetik markörler ve kullanım alanları. *Alatarım* 4(2):27–37
- Han Y, Chagne D, Gasic K, Rikkerink EHA, Beever JE, Gardiner SE, Korban SS (2009) BAC-end sequence-based SNPs and Bin mapping for rapid integration of physical and genetic maps in apple. *Genomics* 93:282–288. <https://doi.org/10.1016/j.ygeno.2008.11.005>
- Han Y, Zheng D, Vimolmangkang S, Khan MA, Beever JE, Korban SS (2011) Integration of physical and genetic maps in apple confirms whole-genome and segmental duplications in the apple genome. *J Exp Bot* 62(14):5117–5130. <https://doi.org/10.1093/jxb/err215>
- Hemmat M, Weeden NF, Brown SK (2003) Mapping and evaluation of *Malus × domestica* microsatellites in apple and pear. *J Am Soc Hortic Sci* 128:515–520
- Hokanson SC, Szewc-McFadden AK, Lamboy WF, McFerson JK (1998) Microsatellite (SSR) markers reveal genetic identities, genetic diversity and relationships in a *Malus × domestica* Borkh. core subset collection. *Theor Appl Genet* 97:671–683. <https://doi.org/10.1007/s001220050943>
- Ikhsan AS, Topcu 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. <https://doi.org/10.1016/j.scienta.2016.10.006>
- Inoue E, Matsuki Y, Anzai H, Evans K (2007) Isolation and characterization of microsatellite markers in Japanese pear (*Pyrus pyrifolia* Nakai). *Mol Ecol Notes* 7:445–447. <https://doi.org/10.1111/j.1471-8286.2006.01612.x>
- Kafkas S, Özkan H, Ak BE, Açar I, Atlı HS, Koyuncu S (2006a) Detecting DNA polymorphism and genetic diversity in a wide pistachio germplasm: comparison of AFLP, ISSR and RAPD markers. *J Am Soc Hortic Sci* 131(4):522–529
- Kafkas S, Kaşka N, Wassimi AN, Padulosi S (2006b) Molecular characterisation of Afghan pistachio accessions by amplified fragment length polymorphisms (AFLPs). *J Hortic Sci Biotechnol* 81:864–868
- Khiari S, Boussaid M, Messaoud C (2015) Genetic diversity and population structure in natural populations of Tunisian Azarole (*Crataegus azarolus* L. var. *aronia* L.) assessed by microsatellite markers. *Biochem Syst Ecol* 59:264–270. <https://doi.org/10.1016/j.bse.2015.01.025>
- Liebhart R, Gianfranceschi L, Koller B, Ryder CD, Tarchini R, Van de Weg E, Gessler C (2002) Development and characterisation of 140 new microsatellites in apple (*Malus × domestica* Borkh.). *Mol Breed* 10:217–241. <https://doi.org/10.1023/A:1020525906332>
- Liu K, Muse SV (2005) PowerMarker: integrated analysis environment for genetic marker data. *Bioinformatics* 21:2128–2129
- Lo EY, Stefanovic S, Dickinson TA (2009) Population genetic structure of diploid sexual and polyploid apomictic hawthorns (*Crataegus*; Rosaceae) in the Pacific Northwest. *Mol Ecol* 18:1145–1160. <https://doi.org/10.1111/j.1365-294X.2009.04091.x>
- MirAli N, Al-Odat M, Haider N, Nabulsi I (2011) The genus *Crataegus* L.: an ecological and molecular study. *Russ J Genet* 47(1):26–32. <https://doi.org/10.1134/S1022795410061018>
- Moriya S, Iwanami H, Kotoda N, Haji T, Okada K, Terakami S, Mimida N, Yamamoto T, Abe K (2012) Aligned genetic linkage maps of apple rootstock cultivar ‘JM7’ and *Malus sieboldii* ‘Sanashi 63’ constructed with novel EST-SSRs. *Tree Genet Genomes* 8:709–723. <https://doi.org/10.1007/s11295-011-0458-3>
- Nas MN (2012) Alıcın (*Crataegus* spp.) kültüre alınması: fırsatlar ve güçlükler. I. Ulusal Alıç Çalıştayı, Malatya, s. 3–8
- Oddou-Muratorio S, Aligon C, Decroocq S, Plomion C, Lamant T, Mush-Demesure B (2001) Microsatellite primers for *Sorbus torminalis* and related species. *Mol Ecol Notes* 1:297–299. <https://doi.org/10.1046/j.1471-8278.2001.00116.x>
- Özgen M, Adak S, Söylemezoğlu G, Ulukan H (2000) Bitki genetik kaynaklarının korunma ve kullanımında yeni yaklaşımlar. Türkiye Ziraat Mühendisliği 5. Teknik Kongresi, Ankara, s. 259–284
- Peakall R, Smouse PE (2012) GenAEx 6.5: genetic analysis in excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28(19):2537–2539. <https://doi.org/10.1093/bioinformatics/bts460>
- Potts SM, Han Y, Khan MA, Kushad MM, Rayburn AL, Korban SS (2012) Genetic diversity and characterization of a core collection of *Malus* germplasm using simple sequence repeats (SSRs). *Plant Mol Biol Rep* 30:827–837. <https://doi.org/10.1007/s11105-011-0399-x>
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155(2):945–959
- Rahmani MS, Shabanian N, Khadivi-Khub A, Woeste KE, Badakhshan H, Alikhani L (2015) Population structure and genotypic variation of *Crataegus pontica* inferred by molecular markers. *Gene* 572:123–129. <https://doi.org/10.1016/j.gene.2015.07.001>
- Rajeb B, Messaoud C, Chograni H, Bejaoui A, Boulila A, Rejeb MN, Boussaid M (2010) Genetic diversity in Tunisian *Crataegus azarolus* L. var. *aronia* L. populations assessed using RAPD markers. *Ann For Sci* 67:512. <https://doi.org/10.1051/forest/2010014>
- Rohlf FJ (2009) NTSYSpc: numerical taxonomy system, ver.2.21c. Setauket: Exeter Publishing
- Scheulke M (2000) An economic method for the fluorescent labeling of PCR fragments. *Nat Biotechnol* 18:233–234
- Serce S, Simsek O, Toplu C, Kamiloglu O, Caliskan O, Gunduz K, Ozgen M, Kacar YA (2011) Relationships among *Crataegus* accessions sampled from Hatay, Turkey, as assessed by fruit characteristics and RAPD. *Genet Resour Crop Evol* 58:933–942. <https://doi.org/10.1007/s10722-010-9633-x>
- Silfverberg-Dilworth E, Matasci CL, Van de Weg WE, Van Kaauwen MPW, Walser M, Kodde LP, Soglio V, Gianfranceschi L, Durel CE, Costa F, Yamamoto T, Koller B, Gessler C, Patocchi A

- (2006) Microsatellite markers spanning the apple (*Malus x domestica* Borkh.) genome. *Tree Genet Genomes* 2:202–224. <https://doi.org/10.1007/s11295-006-0045-1>
- Tanksley SD, Young ND, Paterson AH, Bonierbale MW (1989) RFLP mapping in plant breeding: new tools for an old science. *Nat Technol* 7(3):257–264. <https://doi.org/10.1038/nbt0389-257>
- Testolin R, Marrazzo T, Cipriani G, Quarta R, Verde I, Te Dettori M, Pancaldi M, Sansavini S (2000) Microsatellite DNA in peach (*Prunus persica* L. Batsch) and its use in fingerprinting and testing the genetic origin of cultivars. *Genome* 43(3):512–520
- Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kulper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res* 23(21):4407–4414
- Wang YZ, Zhang JH, Ning N, Sun HY, Yang L (2011) Construction and evaluation of a primary core collection of apricot germplasm in China. *Sci Hortic* 128:311–319. <https://doi.org/10.1016/j.scienta.2011.01.025>
- Wang A, Aldwinckle H, Forsline P, Main D, Fazio G, Brown S, Xu K (2012) EST contig-based SSR linkage maps for *Malus x domestica* cv Royal Gala and an apple scab resistant accession of *M. sieversii*, the progenitor species of domestic apple. *Mol Breed* 29:379–397. <https://doi.org/10.1007/s11032-011-9554-1>
- Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphism amplified by arbitrary primers are useful as genetic markers. *Nucleic Acid Res* 18(22):6531–6535
- Yamamoto T, Kimura T, Sawamura Y, Manabe T, Kotobuki K, Hayashi T, Ban Y, Matsuta N (2002) Simple sequence repeats for genetic analysis in pear. *Euphytica* 124:129–137. <https://doi.org/10.1023/A:1015677505602>
- Yilmaz KU, Yanar M, Ercisli S, Sahiner H, Taskin T, Zengin Y (2010) Genetic relationship among some hawthorn (*Crataegus* spp.) species and genotypes. *Biochem Genet* 48:873–878. <https://doi.org/10.1007/s10528-010-9368-6>
- Zaloğlu S, Kafkas S, Doğan Y, Güney M (2015) Development and characterization of SSR markers from pistachio (*Pistacia vera* L.) and their transferability to eight *Pistacia* Species. *Sci Hortic* 189:94–103. <https://doi.org/10.1016/j.scienta.2015.04.006>
- Zhang Y, Dai HY, Zhang QJ, Li H, Zhang ZH (2008) Assessment of genetic relationship in *Crataegus* genus by the apple SSR primers. *J Fruit Sci* 25(4):521–525
- Zietkiewicz E, Rafalski A, Damian L (1994) Genome fingerprinting by simple sequence repeat (SSR) anchored polymerase chain reaction amplification. *Genomics* 20(2):176–183. <https://doi.org/10.1006/geno.1994.1151>