Cellular and Molecular Biology
E-ISSN : 1165-158X / P-ISSN : 0145-5680
www.cellmolbiol.org

Evaluation of genetic diversity using iPBS-SCoT marker methods in native hawthorn genetic resources and species ıdentification by using DNA barcoding method

Halil Ibrahim Sagbas[1*], Sezai Ercisli[1], Murat Aydin[2], Emre İlhan[3], Recep Aydinyurt[3] Ayse Gul Kasapoglu[3], Selman Muslu[3], Yakup Polat[4]

[1] Faculty of Agriculture, Department of Horticulture, Ataturk University, TR-25240, Erzurum, Türkiye
[2] Faculty of Agriculture, Department of Agricultural Biotechnology, Ataturk University, TR-25240, Erzurum, Türkiye
[3] Faculty of Science, Department of Molecular Biology and Genetics, Erzurum Technical University, TR-25050, Erzurum, Türkiye
[4] Department of Horticulture, Faculty of Agriculture, Van Yuzuncu Yil University, 65200, Van, Türkiye

ARTICLE INFO

Original paper

Article history:
Received: April 28, 2023
Accepted: July 04, 2023
Published: October 31, 2023

Keywords:

Crataegus, hawthorn, molecular characterization, iPBS, SCoT, ITS, genetic diversity

ABSTRACT

Hawthorn is an important medicinal plant that spreads around the world and is used in traditional Chinese medicine. Its flowers and leaves contain flavonoids, vitamins, organic acids and essential oils. Its fruit is consumed as fresh and dried and is an important plant for human health. In this study, iPBS (Inter Primer Binding Site) and SCoT (Start Codon Target Polymorphism) markers were used to analyze genetic variation among 101 hawthorn genotypes collected from Çoruh Valley, Türkiye and ITS markers were used for DNA barcoding. Ten iPBS primers were used and a total of 400 alleles were identified from ten iPBS primers with an average of 40 alleles. PIC values ranged from 0.239 (iPBS 2387) to 0.272 (iPBS 2244). Twenty SCoT primers were used and have an average of 50.05 alleles. The PIC values of the primers ranged from 0.251 (SCoT 2) to 0.297 (SCoT 34). For the DNA barcoding study, it was confirmed that the correct region was amplified and sequenced. The genotypes we used in the study matched 14 different accession numbers by searching a BLASTN in the NCBI. NCBI similarity rates of hawthorn genotypes are between 90.83% and 100%. The study emphasizes the genetic diversity of hawthorn grown from seed and the importance of preserving plant genetic resources.

Doi: http://dx.doi.org/10.14715/cmb/2023.69.10.6

Introduction

Plant genetic resources are accepted as the natural wealth of countries. Compared to underground resources, the disposal of plant genetic resources can be much faster, but their recovery is extremely difficult. Countries should take care of their genetic resources by giving importance to, protecting and evaluating them as necessary, otherwise they may have to buy them from outside. Most of all hawthorn species in the world find natural habitats in Türkiye, and some of them are endemic to Türkiye (1). Coruh Valley, is one of the leading regions in the world in terms of the number of endemic plants (2,3). The fact that Türkiye has three different gene centers such as Euro-Siberian, Mediterranean and Iranian-Turanian Flora has provided a rich plant diversity (4-6). Hawthorn is a species belong to the Rosaceae family, which has 150-200 wild plant species in the world and is known to be valuable, but it has not found the value it deserves in Türkiye, so far (7,8). From this point of view, hawthorn is one of the plant species that are bio-smuggled in Türkiye, although it is a fruit specie that does not have up-to-date production data in the FAO and TUIK databases (9). Between 2015 and 2017, 152 taxa were added to the plant biodiversity worldwide and 25 of them are in Türkiye. There are 21 species, 2 subspecies and 2 hybrid plants in 25 taxa (10). While in previous studies, characterization based on morphological characters was carried out, characterization studies based on molecular markers have been carried out in recent years. Since morphological markers are affected by environmental factors, they show limited polymorphism between similar genotypes, and cannot distinguish an individual with a dominant phenotype from an individual with a recessive phenotype, they are not used much today (11-13). Molecular markers have become the most widely used marker techniques in recent years due to their high number, which are not affected by environmental factors, and could be used in any period of plant growth (6,14).

The term transposon was first used by Barbara McClintock (15) in the corn plant. However, since the genome was thought to have a stable structure in those years, the importance of the subject could not be understood for about 40 years. In the studies carried out in the following years, it was understood that 80-90% of the genome in plants with large genomes consists of transposons (16-18). Owing to transposons, which are mobile DNA fragments, it is possible to see base addition/disappearance, or any duplication (19). iPBS (Inter Primer Binding Site), a retrotransposon-based marker, is based on the amplification of reverse transcriptase primer binding sites (20). Unlike methods of isolation of retrotransposons based on conserved protein-coding regions, iPBS primers directly express polymorphisms for retrotransposon regions in the genome. In the genomes studied, iPBS sequences produce a large

* Corresponding author. Email: hibrahimsagbas@gmail.com

number of clear bands due to their high copy numbers (average 15–50 bands from 100 to 5,000 bp in length) and these bands can be scored easily. iPBS markers have some important advantages, such as genome-wide screening due to their spread over a large part of the eukaryotic genome and having universal primers that can be used in a variety of organisms. (21,22).

SCoT (Start Codon Target Polymorphism) is a DNA marker technique that was discovered in 2009, and works based on short conserved regions surrounding the ATG translation start codon in plants, and performs gene-focused targeting. The basis of the technique is the design and synthesis of single DNA primers for short conserved DNA regions, and PCR-based amplification of target regions by means of these primers, and analysis of data as a result of visualization of bands in standard agarose gel electrophoresis. Advantages of the SCoT marker method compared to other methods: 1) it gives more information about biological properties and universality in plants 2) it is similar to SSR or RAPD technique 3) it binds closely with the target gene (23,24). SCoT has been applied extensively for studies of population genetics, genotyping and even species identification in plants (25-27).

In DNA barcoding studies, which is a DNA sequence-based system, the process works as follows: 1) DNA isolation is performed by taking appropriate tissue samples for DNA isolation of the species desired to be identified 2) PCR amplification and DNA sequencing are conducted 3) the obtained sequences are analyzed and validated 4) the DNA sequences of unidentified species are matched with the DNA sequences in DNA barcode databases and saved data is shared (28,29). Several regions in the plastid genome (atpF-atpH, rpoB, rpoC1, rbcL, ycf5, psbA-trnH, trnL, psbK-psbI) and internal transcribed spacer region (ITS1 and ITS2) plant DNA barcodes (30-32).

The aim of this study was to reveal the genetic diversity of seed-propagated 101 hawthorn genetic resources in Coruh Valley, Türkiye by using iPBS and SCoT molecular marker techniques, and to make identification of the hawthorn species with ITS loci and record them to the NCBI database.

Materials and Methods

The material of this study was collected from Coruh Valley, which is located in the northeastern Anatolian region of Türkiye. The coordinates of the hawthorn plants are as follows: 40º 13' 934"- 40º 37' 913" North latitude and 40º 45' 275"- 42º 05' 917" East longitude. Plants are lowest at 1205 m altitude and highest at 2125 m altitude. The names of plant material were given in Table 10.

Before starting the study, the regions were scanned thoroughly for the availability of experimental material. Fresh leaf samples were taken from the selected genotypes in the 1st and 2nd week of May 2022. The samples were transported in the cold chain to the laboratory and stored at -85 °C.

DNA isolation

DNA isolation was applied by modifying the protocol presented in Emami et al., (33). The purity and concentration of the obtained DNAs were determined by a Nanodrop Lite spectrophotometer. The final concentration of all samples was determined for appropriate purity adjusted to 10 ng/µl and stored at -20°C. Quantities of PCR components for SCoT-iPBS primers and DNA barcoding primers were presented in Table 1.

PCR amplification

PCR amplifications were carried out in SensoQuest LabCycler (SensoQuest GmbH, Goettingen, Germany). The first step is initial denaturation at 95 °C for 3 min and the second step was denaturation at 94 °C for 30 s. In the 3rd step, the hold temperature of the relevant primer is applied for 30 seconds and the 4th step is extended at 72 °C for 1 min. 2nd, 3rd and 4th steps were set to 40 cycles. 5th step is the final extension at 72 °C for 10 min. PCR amplification products were separated on 1.5% agarose gels in 1X TBE buffer stained with ethidium bromide (EtBr) and visualized under UV light. Ten iPBS primers and 2 DNA barcoding primers were used in this study report by Kalendar et al., (21) and Cheng et al., (34), respectively. Twenty SCoT primers were used in the reported by Shen et al., (35), Tyagi et al., (36), Mansoory et al., (37) and Mavlyutov et al., (38). The annealing temperature and primer length do not directly affect the PCR amplification of SCoT primers with high reproducibility. Thus, the annealing temperature of SCoT primers was set at 50 °C according to Collard and Mackill (23). iPBS, SCoT and DNA barcoding primers data that were used in the study were presented in Table 2.

Sequencing, bioinformatics and phylogenetic studies

The obtained PCR products were sequenced by Med-Santek Company (İstanbul, Türkiye) using the Sanger sequencing method. Geneious R8 (39) and Mega v7.0 (40) software were used for bioinformatics studies of ITS nucleotide sequences. The decision sequence was recorded in

Table 1. Quantities of PCR components to be used in a sample.

PCR Components (ITS)	Quantity (µl) (ITS)	PCR Components (SCoT-iPBS)	Quantity (µl) (SCoT-iPBS)
DNA (10 ng/µl)	3	DNA (10 ng/µl)	3
Forward Primer (ITS-p5)	0.5	Primer	2
Reverse Primer (ITS-u4)	0.5		
dNTP (10mM)	0.5	dNTP (10mM)	0.5
$MgCl_2$ (25mM)	2	$MgCl_2$ (25mM)	2
10X PCR Buffer	0.1	10X PCR Buffer	2
5 U/µl Taq-DNA polymerase	0.3	5 U/µl Taq-DNA polymerase	0.3
Distilled water	18.1	Distilled water	10.2
Total Volume	**25**	**Total Volume**	**20**

Table 2. iPBS, SCoT and DNA barcoding primers data used in the study.

Primer	Primer Sequences (5'-3')	AT (°C)	Primer	Primer Sequences (5'-3')	AT (°C)
iPBS 2075	CTCATGATGCCA	50.0	**iPBS 2244**	GGAAGGCTCTGATTACCA	53.7
iPBS 2083	CTTCTAGCGCCA	45.7	**iPBS 2387**	GCGCAATACCCA	50.0
iPBS 2219	GAACTTATGCCGATACCA	51.5	**iPBS 2394**	GAGCCTAGGCCA	48.5
iPBS 2222	ACTTGGATGCCGATACCA	55.0	**iPBS 2400**	CCCCTCCTTCTAGCGCCA	57.4
iPBS 2226	CGGTGACCTTTGATACCA	50.0	**iPBS 2415**	CATCGTAGGTGGGCGCCA	50.0
SCoT 1	CAACAATGGCTACCACCA	50.0	**SCoT 14**	ACGACATGGCGACCACGC	50.0
SCoT 2	CAACAATGGCTACCACCC	50.0	**SCoT 15**	ACGACATGGCGACCGCGA	50.0
SCoT 3	CAACAATGGCTACCACCG	50.0	**SCoT 20**	ACCATGGCTACCACCGCG	50.0
SCoT 4	CAACAATGGCTACCACCT	50.0	**SCoT 21**	ACGACATGGCGACCCACA	50.0
SCoT 5	CAACAATGGCTACCACGA	50.0	**SCoT 28**	CCATGGCTACCACCGCCA	50.0
SCoT 8	CAACAATGGCTACCACGT	50.0	**SCoT 30**	CCATGGCTACCACCGGCG	50.0
SCoT 9	CAACAATGGCTACCAGCA	50.0	**SCoT 33**	CCATGGCTACCACCGCAG	50.0
SCoT 11	AAGCAATGGCTACCACCA	50.0	**SCoT 34**	ACCATGGCTACCACCGCA	50.0
SCoT 12	ACGACATGGCGACCAACG	50.0	**SCoT 60**	ACAATGGCTACCACCACA	50.0
SCoT 13	ACGACATGGCGACCATCG	50.0	**SCoT 72**	CCATGGCTACCACCGCCC	50.0
ITS-p5	CCTTATCAYTTAGAGGAAGGAG	55.0	**ITS-u4**	RGTTTCTTTTCCTCCGCTTA	55.0

AT: annealing temperature

the GenBank database by obtaining the accession number. BLAST (Basic Local Alignment Search Tool) has been performed on the NCBI nucleotide database to confirm that the correct region has been amplified and sequenced. The ITS1, 5.8S, and ITS2 regions of the DNA sequence were identified and labeled by comparison with homologous sequences. The repeat regions in the ITS sequence of the *Crataegus* spp. were calculated with Phobos 3.3 (41) software.

For phylogenetic studies, hawthorn sequences with 90% and above similarity in BLAST results were downloaded from the database. Alignment of the sequences was done with the ClustalW algorithm (42).

Evaluation of molecular data

Genetic similarities were calculated using the Dice similarity index (43). The PIC value for each amplified loci was calculated using Power Marker 3.0. NTSYS-PC package software was used in cluster analysis and dendrograms were drawn according to UPGMA method. (Na) Number of distinct alleles, (Ne) effective number of alleles, (h) gene diversity (44) and (I) Shannon's information index (45) were carried out using POPGEN32 software (v3.2 Microsoft Windows-Based Freeware for Population Genetics Analysis) (46). The genetic structure of the hawthorns was determined using a model based on the clustering algorithm (STRUCTURE v.2.2), which genetically separates groups according to allele frequency (47). Molecular variance analysis was performed using the AMOVA method in the GenAlex program (48) in order to determine the within-among population variation ratio.

Results

Evaluation of genetic diversity with the iPBS marker method

In our study, a total of 400 alleles were detected according to the data obtained from the iPBS primers, and the average number of alleles per locus was shown to be 40. The locus that produced the most alleles was iPBS 2075 with 49 alleles, while the locus that produced the least alleles was iPBS 2222 with 22 alleles (Table 3). The PIC value of the iPBS markers ranged from 0.239 (iPBS 2387) to 0.272 (iPBS 2244) and was calculated to be an average of 0.258. The number of effective alleles has been reported to range from 1.353 (iPBS 2400) to 1.541 (iPBS 2222), with an average of 1.420. According to the Shannon information index, the index values of the markers ranged from 0.366 (iPBS 2400) to 0.480 (iPBS 2387), with an average of 0.418. The lowest gene diversity was observed in the iPBS 2400 marker with 0.230; however, the highest gene diversity was observed in the iPBS 2222 marker with 0.316. The mean gene diversity value of the loci was detected to be 0.268 (Table 3). The gel image of iPBS 2226 locus is presented in Figure 1.

According to the results of the clustering analysis, the genotype pairs with the lowest similarity were 25C10-25C51 (0.60), the genotype pairs with the highest genetic similarity were 25C07-25C74 (0.99). The percentage of polymorphism between genotypes was quite high. The similarity coefficient was shown between 0.70-0.99, and was presented in Figure 3. The percentages of genetic diversity explained by the three main axes of principal coordinate analysis (PCoA) are 10.70, 7.78 and 6.06, respectively. The first three components explained 24.54% of the genetic diversity, the first two principal components explained 18.48% of the genetic diversity and the first principal component explained 10.70% of the genetic diversity (Table 4).

When the similarity matrix was examined, it was seen that the closest population pair to each other was Oltu-Tortum (0.987). The other closest population pairs were

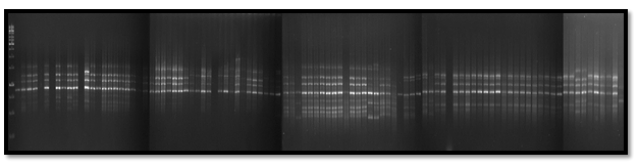

Figure 1. Gel image of iPBS 2226 primer.

Table 3. Data from iPBS primers.

	Primer Name	N	Na	Ne	I	h	Allele Number	Number of Bands	Bands Per Allele	PIC
1	iPBS 2075	101	2	1.396	0.386	0.246	49	945	19	0.259
2	iPBS 2083	101	2	1.451	0.441	0.282	47	936	20	0.256
3	iPBS 2219	101	2	1.414	0.410	0.260	35	658	19	0.261
4	iPBS 2222	101	2	1.541	0.474	0.316	22	527	24	0.241
5	iPBS 2226	101	2	1.420	0.417	0.266	37	696	19	0.260
6	iPBS 2244	101	2	1.362	0.376	0.235	46	726	16	0.272
7	iPBS 2387	101	2	1.529	0.480	0.315	38	934	25	0.239
8	iPBS 2394	101	2	1.448	0.437	0.281	44	855	19	0.258
9	iPBS 2400	101	2	1.353	0.366	0.230	48	777	16	0.270
10	iPBS 2415	101	2	1.383	0.397	0.250	34	577	17	0.267
	TOTAL						**400**			
	AVERAGE			**1.420**	**0.418**	**0.268**	**40**			**0.258**

N = Number of Genotypes, Na = Number of distinct alleles, Ne = Effective number of alleles, I = Shannon's information index, h = Gene diversity, PIC = Polymorphism information index.

Table 4. Percentage of variation obtained with iPBS markers.

Components	1	2	3
Percentage of Variance (%)	10.70	7.78	6.06
Cumulative Analysis of Variance (%)	10.70	18.48	24.54

Table 5. Pairwise population matrix of Nei genetic distance (iPBS).

Tortum	Uzundere	Oltu	İspir	Pazaryolu	
1.000					Tortum
0.981	1.000				Uzundere
0.987	0.973	1.000			Oltu
0.984	0.969	0.986	1.000		İspir
0.975	0.963	0.979	0.973	1.000	Pazaryolu

shown to be Oltu-İspir (0.986) and Tortum-İspir (0.984). It has been determined that the Uzundere-Pazaryolu population pair is the most distant population with 0.963 (49) genetic distance (Table 5).

According to the results of molecular analysis of variance (AMOVA), it was observed that 91% of the variation was within the population and 9% was between the populations. According to the results of the genetic structure analysis of hawthorn genotypes based on iPBS markers, hawthorn genotypes were divided into two subpopulations (ΔK=2) (Figure 2). The significance of the variance components was tested with 100,000 random permutations. The red and green colors represent a different population group in the estimated group plot. Subpopulation 1 and 2 is shown in red and green, respectively (Figure 4).

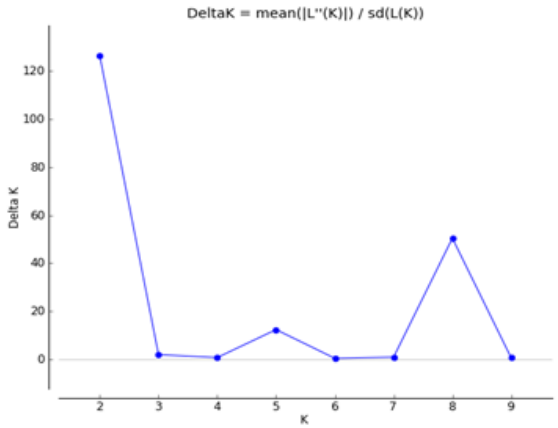

Figure 2. Graph of K (ΔK) values of hawthorn genotypes obtained by using the STRUCTURE program based on iPBS markers.

Evaluation of genetic diversity with the SCoT marker method

In our study, a total of 1001 alleles were detected according to the data obtained from the SCoT primers, and the average number of alleles per locus was found to be 50.05. In the study, the locus with the most alleles was the primer SCoT 13 with 60 alleles, while the locus with the least alleles was the primer SCoT 2 with 37 alleles. It was calculated that the PIC value of the loci used in the study ranged from 0.251 (SCoT 2) to 0.297 (SCoT 34), with an average of 0.269 (Table 6). The number of effective alleles has been reported to range from 1.287 (SCoT 34) to 1.530 (SCoT 60), with an average of 1.451. Shannon information index values of the markers ranged from 0.314 (SCoT 34) to 0.489 (SCoT 33), with an average of 0.438. The lowest gene diversity was observed in the SCoT 34 marker with 0.190; however, the highest gene diversity was observed in the SCoT 33 marker with 0.321. The mean gene diversity value of the loci was found to be 0.282 (Table 6).

It was shown that the genotype pairs with the lowest similarity were 25C08-25C61 (0.61), but the genotype pairs with the highest genetic similarity were 25C74-25C76 (0.87) according to cluster analysis results obtained with SCoT markers. The percentage of polymorphism between

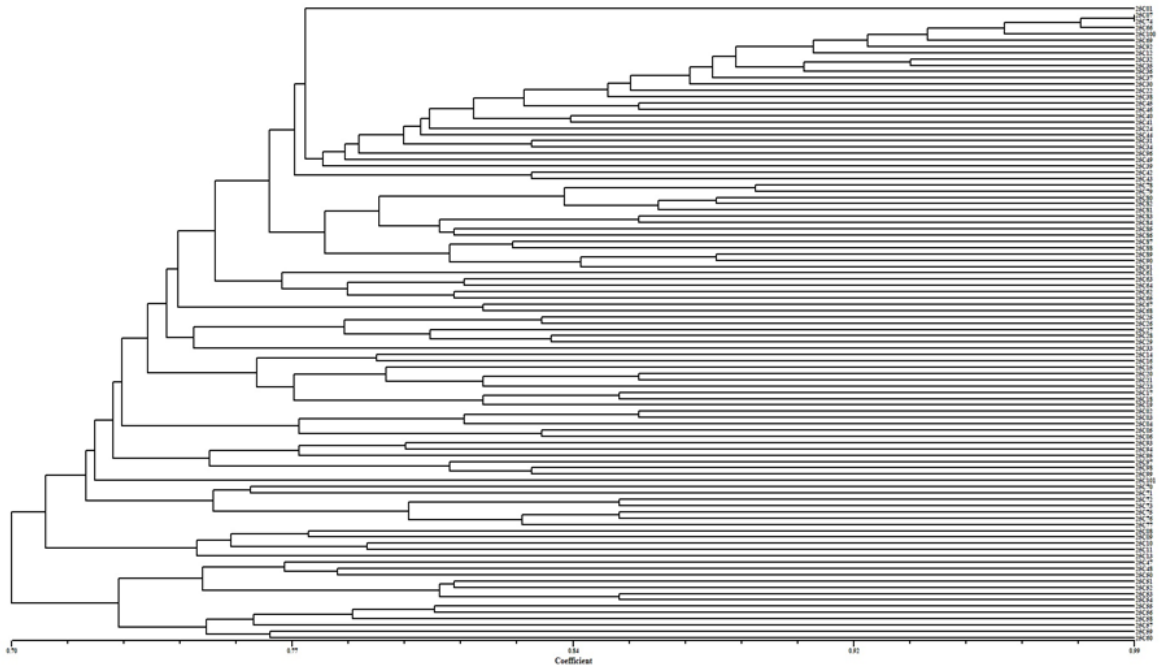

Figure 3. Cluster analysis graph created by UPGMA method using iPBS markers.

Table 6. Data from SCoT primers.

	Loci	N	Na	Ne	I	h	Allele Number	Number of Bands	Bands Per Allele	PIC
1	SCoT 1	101	2	1.472	0.447	0.289	54	1155	21	0.265
2	SCoT 2	101	2	1.518	0.460	0.305	37	925	25	0.251
3	SCoT 3	101	2	1.445	0.423	0.272	41	816	20	0.271
4	SCoT 4	101	2	1.375	0.388	0.243	50	819	16	0.284
5	SCoT 5	101	2	1.360	0.388	0.239	50	796	16	0.286
6	SCoT 8	101	2	1.448	0.447	0.286	58	1142	20	0.272
7	SCoT 9	101	2	1.449	0.425	0.274	50	992	20	0.271
8	SCoT 11	101	2	1.477	0.461	0.298	58	1271	22	0.263
9	SCoT 12	101	2	1.401	0.428	0.268	57	999	18	0.280
10	SCoT 13	101	2	1.503	0.482	0.314	60	1316	22	0.263
11	SCoT 14	101	2	1.477	0.468	0.302	53	1132	21	0.265
12	SCoT 15	101	2	1.435	0.433	0.276	53	1015	19	0.274
13	SCoT 20	101	2	1.506	0.469	0.307	48	1155	24	0.255
14	SCoT 21	101	2	1.484	0.437	0.287	47	1062	23	0.260
15	SCoT 28	101	2	1.416	0.422	0.269	51	954	19	0.275
16	SCoT 30	101	2	1.480	0.466	0.301	54	1156	21	0.265
17	SCoT 33	101	2	1.529	0.489	0.321	45	1052	23	0.258
18	SCoT 34	101	2	1.287	0.314	0.190	38	482	13	0.297
19	SCoT 60	101	2	1.530	0.476	0.315	45	1079	24	0.255
20	SCoT 72	101	2	1.426	0.430	0.273	52	969	19	0.276
	TOTAL						**1,001**			
	AVERAGE			**1.451**	**0.438**	**0.282**	**50.05**			**0.269**

N = Number of Genotypes. Na = Number of distinct alleles. Ne = Effective number of alleles. I = Shannon's information index. h = Gene diversity. PIC = Polymorphism information index.

the genotypes was quite high. The similarity coefficients were shown between 0.69-0.87, and were presented in Figure 6. The percentages of genetic diversity explained by the three main axes of principal coordinate analysis (PCoA) are 6.11, 5.31 and 4.03, respectively. The first three components explained 15.45% of the genetic diversity, the first two principal components explained 11.42% of the genetic diversity, and the first principal component explained 6.11% of the genetic diversity (Table 7).

When the similarity matrix was examined, it was seen that the closest population pair to each other was Oltu-Tortum (0.980). The other closest population pairs were Oltu-İspir (0.978) and Tortum-İspir (0.975). The Uzundere-Pazaryolu population pair is the most distant population with

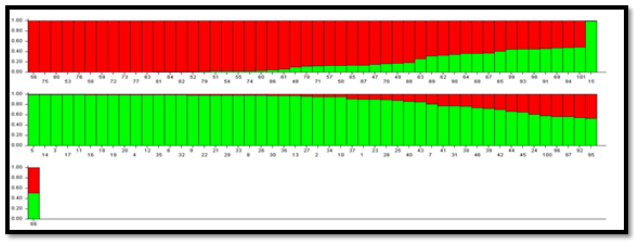

Figure 4. Graphical representation of genetic structure analysis of the genotypes based on iPBS markers.

0.943 (49) genetic distance (Table 8).

According to the results of molecular analysis of variance (AMOVA) performed with the SCoT marker data of the genotypes in our study, it was observed that 88% of the variation was within the population and 12% was between the populations. According to the results of genetic structure analysis based on SCoT markers, our hawthorn genotypes were distributed three sub-populations (ΔK=3) (Figure 5). The red, green and blue colors represent different population groups in the graph (Figure 7). According to the data in the Google Earth application, the coordinates of districts are following: Tortum is 40º17'40.72" (N)/41º32'59.81" (E), Uzundere is 40º36'41.64" (N)/41º37'42.52" (E), Oltu is 40º32'45.58" (N)/ 41º59'44.26" (E), İspir is 40º29'03.53" (N)/ 41º00'01.36" (E) and Pazaryolu is 40º26'45.51" (N)/ 40º43'05.50" (E).

Evaluation of genetic diversity by DNA barcoding method

The genotypes matched 14 different accession numbers registered in the NCBI database (Table 9). Sequence data of 9 genotypes (25C18, 25C32, 25C37, 25C38, 25C62, 2569, 25C74, 25C77, and 25C92) could not be obtained in this sequencing study (Table 10). However, sequences in the range of 606-662 bp were obtained for 92 hawthorn genotypes. NCBI similarity rates were detected to range from 90.83% (25C66 / *C. tanacetifolia* - *C. maximowiczii voucher*) to 100% (25C63 / *Crataegus meyeri* - *Crataegus laevigata*), and phylogenetic tree has been created. The cluster analysis plot of the sequence study is presented in Figure 8. Galaxy Gala apple cultivar is placed as outgroup in the dendrogram.

Discussion

The iPBS markers developed by Kalendar et al., (21) are called "universal retrotransposon markers" and they can be used in all eukaryotic organisms. They have become widely used in genetic diversity studies due to their general and simple applicability (50,51). Gurlen et al., (52) carried out molecular characterization studies using the iPBS marker method on 25 hawthorn genotypes taken from the province of Bolu in the western black sea region of Türkiye. 68 bands were obtained by amplifying the 6 iPBS primers used and 65 of these bands (95.59%) showed polymorphism. It was also stated that the PIC values were between 0.12 and 0.42. Ferrazzini et al., (53) conducted a study on 6 different populations of the *C. monogyna* specie in Italy to determine the molecular variance ratio and used the RAPD marker method in the study. They found that 79.75% of the molecular variance was within the population and 20.25% between the populations. Erfani-Moghadam et al., (54) conducted genetic diversity research on four different species of *Crataegus* genus, which is one of the important forest plants of Iran, by principal component analysis (PCA). 85% of the diversity consists of the first five components and the first two components correspond to approximately 55.24% of the entire variability. Rahmani et al., (55) conducted a study to detect the polymorphism levels of genotypes using SCoT, ISSR and IRAP markers in 164 genotypes from 14 different populations of hawthorn (*C. pontica*). The observed allele counts are 126 (IRAP), 254 (ISSR), and 199 (SCoT). Polymorphism rates are 90.48% (IRAP), 93.37% (ISSR) and 83.78% (SCoT). PIC values range from 0.16-0.28 with SCoT marker and mean 0.23; 0.14-0.30 with ISSR marker and mean 0.23;

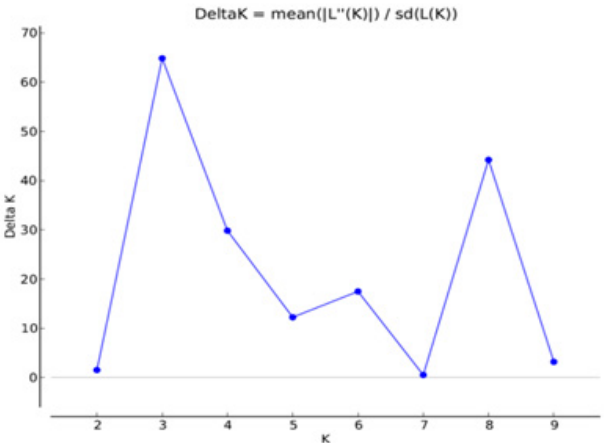

Figure 5. Graph of K (ΔK) values of hawthorn genotypes obtained by using the STRUCTURE program based on SCoT markers.

Table 7. Percentage of variation obtained with SCoT markers.

Components	1	2	3
Percentage of Variance (%)	6.11	5.31	4.03
Cumulative Analysis of Variance (%)	6.11	11.42	15.45

Table 8. Pairwise population matrix of Nei genetic distance (SCoT).

Tortum	Uzundere	Oltu	İspir	Pazaryolu	
1.000					Tortum
0.963	1.000				Uzundere
0.980	0.966	1.000			Oltu
0.975	0.955	0.978	1.000		İspir
0.966	0.943	0.964	0.964	1.000	Pazaryolu

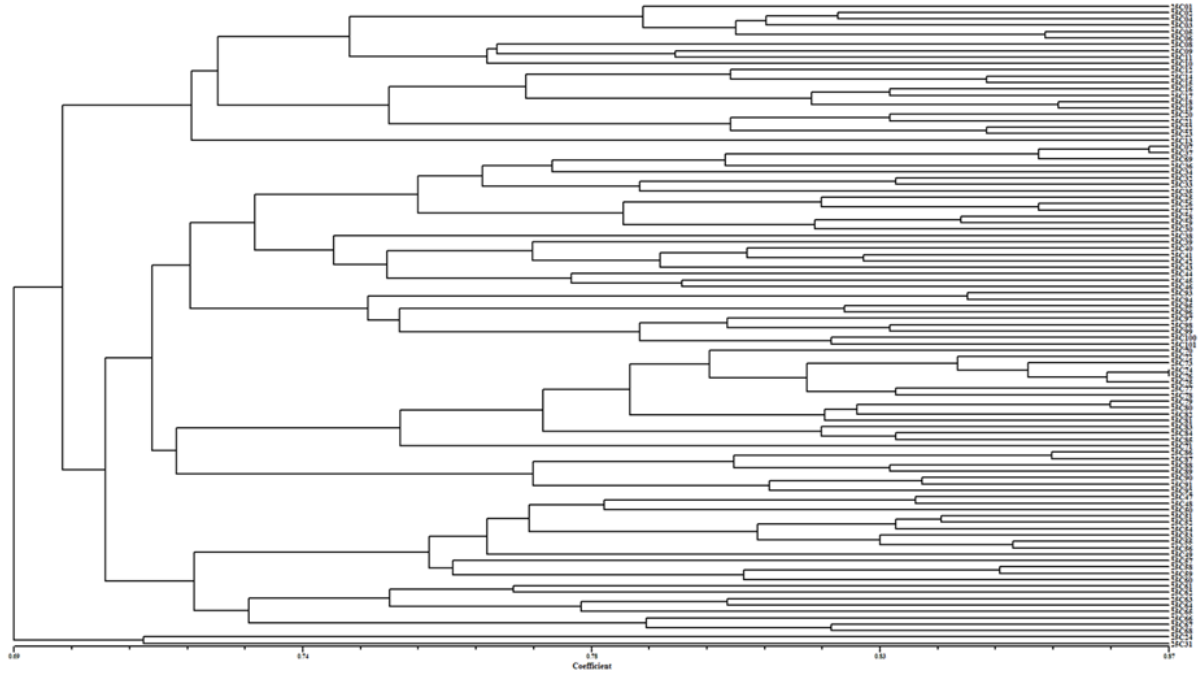

Figure 6. Cluster analysis graph created by UPGMA method using SCoT markers.

Table 9. NCBI accession numbers and species names that match hawthorn genotypes.

	Accession Numbers	NCBI Species Names		Accession Numbers	NCBI Species Names
1	EF127015.1	*C. laevigata voucher*	**8**	EF127016.1	*C. heldreichii*
2	EF127017.1	*Crataegus. sp. EYYL-2006*	**9**	EF127038.1	*C. hupehensis voucher*
3	EU500465.1	*C. heldreichii voucher*	**10**	EU500466.1	*C. laevigata voucher*
4	EU785942.1	*C. laevigata voucher*	**11**	KJ506855.1	*C. laevigata*
5	MN722068.1	*C. hupehensis voucher*	**12**	MT113325.1	*C. azarolus*
6	MW362361.1	*C. monogyna*	**13**	MZ686467.1	*C. maximowiczii voucher*
7	MZ686480.1	*C. pinnatifida var. major*	**14**	MZ686490.1	*C. altaica*

0.16-0.27 with IRAP marker and mean 0.24. Emami et al., (33) carried out a study with SCoT primers on 201 genotypes of *C. aronia, C. ambigua, C. pseudoheterophylla, C. atrosanguinea, C. orientalis* and *C. pontica* in Iran. According to the Jaccard similarity index, the genetic distance between genotypes was found to be in the range of 0.02-0.62. It has been stated that the SCoT marker method can be used to determine the genetic relationships of hawthorn genetic resources. A group of researchers conducted a genetic diversity study on 23 hawthorn populations in Northern Ireland. Nuclear and chloroplast SSR markers were used in the study. Molecular variance analysis with nuclear DNA SSR primers revealed that 99.08% of the variation was within the population and 0.92% was between the populations. In the molecular variance analysis performed with microsatellite markers with chloroplast DNA, it was stated that 96.15% of the variation was within the population and 3.85% was between the populations. According to the results of the study, the genetic diversity was high and at a comparable level (56). Yildiz et al., (57) carried out a study to evaluate the molecular diversity of 22 hawthorn genotypes in Kayseri (Turkey) by using ISSR markers. 13 primers were used in the study, and the polymorphism rate was 75.24%. In the dendrogram created, the similarity rate of the genotypes was found to be in the range of 0.71-0.88. Zhang et al., (58) conducted a study using SCoT markers to determine the genetic diversity of 36 hawthorn genotypes in China. They obtained

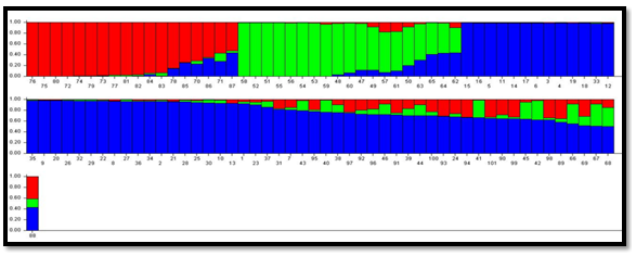

Figure 7. Graphical representation of genetic structure analysis of the genotypes based on SCoT markers.

148 polymorphic bands in the study, and stated that SCoT primers can be used in genetic studies in hawthorn species. Zarrei et al., (59) conducted sequence analysis using four different DNA barcodes (rbcL, matK, psbA-trnH and ITS2) on 355 genotypes of 93 hawthorn species in Canada. It was stated that the barcode primers used were not sufficient to explain the phylogenetic relationships of hawthorn genotypes. They attributed this to the absence of variation in the plastid loci (rbcL, matK, psbA-trnH) used in the study. In addition, it was concluded that the internal transcribed spacer region (ITS2) into the nuclear gene did not give a clear result when faced with the high frequency of allopolyploidy in hawthorns. Shin et al., (60) conducted a DNA barcoding study in hawthorn species. ITS, trnH, psbA and rbcL barcodes used in the study could not sufficiently differentiate between species.

Table 10. NCBI data of hawthorn genotypes.

Genotypes	Species Names	NCBI Species Names	Similarity Rates (%)	NCBI Registration Numbers
25C01	*C. orientalis subsp. orientalis*	*C. azarolus*	99.68	OP503222
25C02	*C. orientalis subsp. orientalis*	*C. azarolus*	99.84	OP503223
25C03	*C. orientalis subsp. orientalis*	*C. azarolus*	99.84	OP503224
25C04	*C. orientalis subsp. orientalis*	*C. azarolus*	99.68	OP503225
25C05	*C. orientalis subsp. orientalis*	*C. azarolus*	99.68	OP503226
25C06	*C. orientalis subsp. orientalis*	*C. azarolus*	98.88	OP503227
25C07	*C. meyeri Pojark.*	*C. azarolus*	99.68	OP503197
25C08	*C. meyeri Pojark.*	*C. laevigata voucher*	99.53	OP503198
25C09	*C. orientalis subsp. orientalis*	*C. azarolus*	99.05	OP503228
25C10	*C. meyeri Pojark.*	*C. azarolus*	99.05	OP503199
25C11	*C. meyeri Pojark.*	*C. azarolus*	99.36	OP503200
25C12	*C. meyeri Pojark.*	*C. laevigata voucher*	99.68	OP503201
25C13	*C.* x *bornmülleri* ZABEL	*C. laevigata voucher*	97.14	OP503272
25C14	*C. meyeri Pojark.*	*C. monogyna*	99.53	OP503182
25C15	*C. orientalis* subsp. *orientalis*	*C. monogyna*	99.84	OP503204
25C16	*C. orientalis subsp. orientalis*	*C. monogyna*	99.68	OP503205
25C17	*C. orientalis subsp. orientalis*	*C. monogyna*	99.53	OP503206
25C18	*C. orientalis subsp. orientalis*	-	-	-
25C19	*C. orientalis subsp. orientalis*	*C. laevigata*	99.68	OP503207
25C20	*C. meyeri Pojark.*	*C. laevigata*	99.05	OP503183
25C21	*C. pentagyna* WALDST. ET KIT. EX WILLD.	*C. monogyna*	99.21	OP503230
25C22	*C. meyeri Pojark.*	*C. monogyna*	99.37	OP503184
25C23	*C. pentagyna* WALDST. ET KIT. EX WILLD.	*C. monogyna*	99.53	OP503231
25C24	*C. orientalis subsp. orientalis*	*C. azarolus*	99.53	OP503208
25C25	*C. orientalis subsp. orientalis*	*C. azarolus*	99.69	OP503209
25C26	*C. pentagyna* WALDST. ET KIT. EX WILLD.	*C. azarolus*	99.53	OP503232
25C27	*C. pentagyna* WALDST. ET KIT. EX WILLD.	*C. laevigata*	99.68	OP503233
25C28	*C. meyeri Pojark.*	*C. azarolus*	99.22	OP503185
25C29	*C. orientalis subsp. orientalis*	*C. heldreichii voucher*	99.37	OP503210
25C30	*C. meyeri Pojark.*	*C. azarolus*	98.89	OP503186
25C31	*C. orientalis subsp. orientalis*	*C. monogyna*	99.84	OP503211
25C32	*C. oxyacantha Linn.*	-	-	-
25C33	*C. oxyacantha Linn.*	*C. monogyna*	99.53	OP503229
25C34	*C. monogyna* Jacq. var *monogyna*	*C. laevigata voucher*	95.86	OP503202
25C35	*C. orientalis subsp. orientalis*	*C. monogyna*	99.68	OP503212
25C36	*C.* x *bornmülleri* ZABEL	*C. laevigata voucher*	95.38	OP503273
25C37	*C. oxyacantha Linn.*	-	-	-
25C38	*C. oxyacantha Linn.*	-	-	-
25C39	*C. monogyna* Jacq. var *monogyna*	*C. hupehensis voucher*	96.39	OP503203
25C40	*C. meyeri Pojark.*	*C. monogyna*	99.21	OP503187
25C41	*C. meyeri Pojark.*	*C. monogyna*	98.42	OP503188
25C42	*C. orientalis subsp. orientalis*	*C. azarolus*	99.21	OP503213
25C43	*C. orientalis subsp. orientalis*	*C. monogyna*	99.05	OP503214
25C44	*C. tanacetifolia (Lam.) Pers.*	*C. monogyna*	94.18	OP503241
25C45	*C. orientalis subsp. orientalis*	*C. laevigata*	99.06	OP503217
25C46	*C. pentagyna* WALDST. ET KIT. EX WILLD.	*C. monogyna*	99.84	OP503234
25C47	*C. orientalis subsp. orientalis*	*C. monogyna*	98.58	OP503216
25C48	*C. orientalis subsp. orientalis*	*C. laevigata*	99.22	OP503218
25C49	*C. meyeri Pojark.*	*C. monogyna*	99.53	OP503189
25C50	*C. pentagyna* WALDST. ET KIT. EX WILLD.	*C. monogyna*	99.68	OP503235
25C51	*C. meyeri Pojark.*	*C. laevigata voucher*	98.24	OP503190
25C52	*C. orientalis subsp. orientalis*	*C. azarolus*	99.38	OP503215
25C53	*C. meyeri Pojark.*	*C. laevigata*	99.69	OP503191
25C54	*C. meyeri Pojark.*	*C. laevigata*	98.59	OP503192
25C55	*C. meyeri Pojark.*	*C. monogyna*	99.68	OP503193
25C56	*C. meyeri Pojark.*	*C. monogyna*	99.52	OP503194
25C57	*C. pentagyna* WALDST. ET KIT. EX WILLD.	*C. laevigata*	99.08	OP503236
25C58	*C. orientalis subsp. orientalis*	*C. monogyna*	98.09	OP503219
25C59	*C. orientalis subsp. orientalis*	*C. azarolus*	99.54	OP503220
25C60	*C. orientalis subsp. orientalis*	*C. laevigata*	99.37	OP503221
25C61	*C. pentagyna* WALDST. ET KIT. EX WILLD.	*C. heldreichii*	99.84	OP503237
25C62	*C. pentagyna* WALDST. ET KIT. EX WILLD.	-	-	-
25C63	*C. meyeri Pojark.*	*C. laevigata*	100	OP503195

25C64	*C. meyeri Pojark.*	*C. laevigata*	99.53	OP503196
25C65	*C. pentagyna* WALDST. ET KIT. EX WILLD.	*C. laevigata voucher*	96.83	OP503238
25C66	*C. tanacetifolia (Lam.) Pers.*	*C. maximowiczii voucher*	90.83	OP503242
25C67	*C. tanacetifolia (Lam.) Pers.*	*C. laevigata voucher*	94.79	OP503243
25C68	*C. tanacetifolia (Lam.) Pers.*	*C. laevigata voucher*	93.08	OP503244
25C69	*C. tanacetifolia (Lam.) Pers.*	-	-	-
25C70	*C. tanacetifolia (Lam.) Pers.*	*C. laevigata*	92.88	OP503245
25C71	*C. pentagyna* WALDST. ET KIT. EX WILLD.	*C. heldreichii*	98.73	OP503239
25C72	*C. tanacetifolia (Lam.) Pers.*	*C. hupehensis*	94.44	OP503246
25C73	*C. tanacetifolia (Lam.) Pers.*	*C. laevigata voucher*	95.71	OP503247
25C74	*C. tanacetifolia (Lam.) Pers.*	-	-	-
25C75	*C. tanacetifolia (Lam.) Pers.*	*C. altaica*	95.63	OP503248
25C76	*C. tanacetifolia (Lam.) Pers.*	*C. laevigata voucher*	95.58	OP503249
25C77	*C. tanacetifolia (Lam.) Pers.*	-	-	-
25C78	*C. pentagyna* WALDST. ET KIT. EX WILLD.	*C. laevigata voucher*	96.99	OP503240
25C79	*C. tanacetifolia (Lam.) Pers.*	*C. laevigata voucher*	96.81	OP503250
25C80	*C. tanacetifolia (Lam.) Pers.*	*C. pinnatifida var. major*	96.46	OP503251
25C81	*C. tanacetifolia (Lam.) Pers.*	*C. laevigata voucher*	95.40	OP503252
25C82	*C. tanacetifolia (Lam.) Pers.*	*C. azarolus*	92.26	OP503253
25C83	*C. tanacetifolia (Lam.) Pers.*	*C. sp. EYYL-2006*	96.58	OP503254
25C84	*C. tanacetifolia (Lam.) Pers.*	*C. hupehensis voucher*	95.29	OP503255
25C85	*C. tanacetifolia (Lam.) Pers.*	*C. laevigata voucher*	94.79	OP503256
25C86	*C. tanacetifolia (Lam.) Pers.*	*C. hupehensis voucher*	95.25	OP503257
25C87	*C. tanacetifolia (Lam.) Pers.*	*C. laevigata voucher*	95.09	OP503258
25C88	*C. tanacetifolia (Lam.) Pers.*	*C. laevigata voucher*	95.40	OP503259
25C89	*C. tanacetifolia (Lam.) Pers.*	*C. laevigata voucher*	95.09	OP503260
25C90	*C. tanacetifolia (Lam.) Pers.*	*C. laevigata voucher*	95.09	OP503261
25C91	*C. tanacetifolia (Lam.) Pers.*	*C. hupehensis voucher*	95.73	OP503262
25C92	*C. tanacetifolia (Lam.) Pers.*	-	-	-
25C93	*C. tanacetifolia (Lam.) Pers.*	*C. azarolus*	99.36	OP503263
25C94	*C. tanacetifolia (Lam.) Pers.*	*C. altaica*	96.60	OP503264
25C95	*C. tanacetifolia (Lam.) Pers.*	*C. heldreichii*	98.89	OP503265
25C96	*C. tanacetifolia (Lam.) Pers.*	*C. laevigata*	99.68	OP503266
25C97	*C. tanacetifolia (Lam.) Pers.*	*C. laevigata voucher*	96.62	OP503267
25C98	*C. tanacetifolia (Lam.) Pers.*	*C. laevigata voucher*	96.78	OP503268
25C99	*C. tanacetifolia (Lam.) Pers.*	*C. pinnatifida var. major*	96.27	OP503269
25C100	*C. tanacetifolia (Lam.) Pers.*	*C. maximowiczii*	91.73	OP503270
25C101	*C. tanacetifolia (Lam.) Pers.*	*C. laevigata voucher*	96.85	OP503271

Previous studies conducted on different wild edible fruits showed that due to rich gene combination, those plants showed high genetic diversity and adaption (61-66).

DNA barcoding studies of the seed-propagated hawthorn (*Crataegus* spp.) population, which are densely found in Erzurum province in the Coruh Valley of Türkiye, (which is considered one of the most important plant genetic resource centers in the world), has not been carried out until today. In addition, a limited number of genetic diversity studies on hawthorn species with SCoT and iPBS markers in the world and in Türkiye reveals the original value of the study. As a result, the genetic diversity of the hawthorn population propagated from seed in the region was investigated, and available species were registered in the NCBI database, and it was ensured that putative new hawthorn cultivar candidates could be revealed. In the first stage of this study, hawthorn genotypes that are morphologically different from each other as much as possible were selected. In these dendrograms created by molecular characterization studies, it was observed that the genotypes were not clearly separated from each other morphologically. Individuals who look morphologically similar may have very different characters from each other at the molecular level. Thus, the importance of conducting studies at the DNA level has been demonstrated once again.

In the dendrogram and genetic structure analysis formed with iPBS primers genotypes are clustered in 2 main groups. The similarity coefficients of the genotypes were detected between 0.70-0.99. According to the results of molecular analysis of variance (AMOVA), it was observed that 91% of the variation was within the population and 9% was among the populations.

According to the SCoT marker data, genotypes were shown in 2 main groups in the phylogenetic tree drawn. The similarity coefficients in the dendrogram were shown to be between 0.69-0.87. Hawthorn genotypes are placed in three sub-populations in genetic structure analysis. It was observed that 88% of the variation was within the population and 12% was among the populations into the molecular analysis of variance.

According to the results of our study, it is understood that iPBS and SCoT marker methods can be easily used in genetic diversity studies in hawthorn species. However, it was observed that the SCoT marker method gave more descriptive and clear results compared to the iPBS marker method. Moreover, it is clearly stated in the previous studies that the high level of polyploidy specific to the hawthorn species is a major problem in DNA barcoding studies. Primers iPBS 2244, iPBS 2400, SCoT 34 and SCoT 72 with high PIC values can be preferred in studies on the expression of genetic diversity in hawthorn (*Crataegus* spp.).

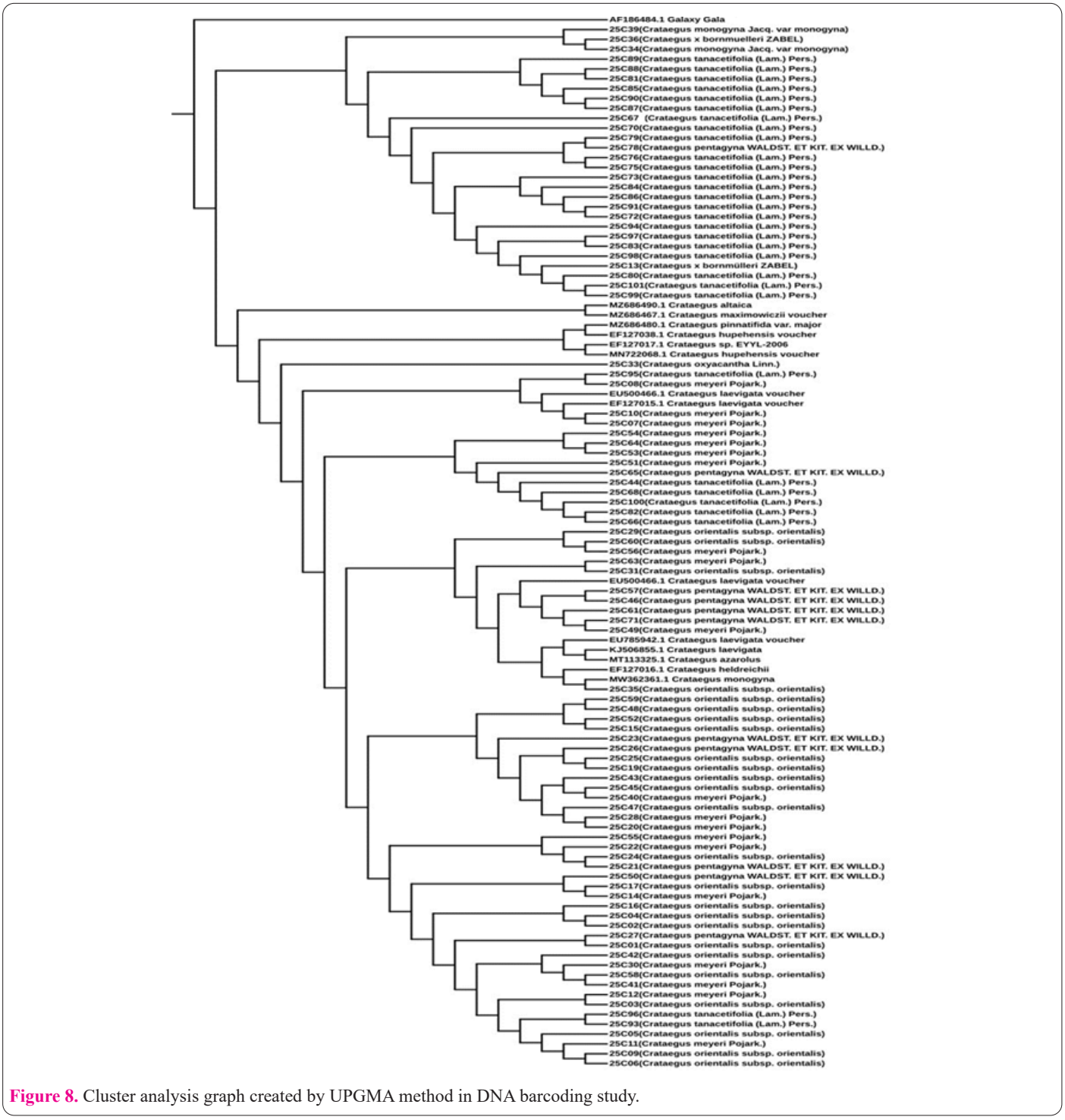

Figure 8. Cluster analysis graph created by UPGMA method in DNA barcoding study.

DNA barcode studies in plant species will inevitably make great contributions to taxonomy and genetic diversity studies. The number of plant DNA barcode studies to be performed on species such as hawthorn, which are particularly taxonomically problematic, should be increased. After the barcode information of the species is officially created, species identification should be made, and the species should be recorded in databases. More DNA barcoding studies with the use of new plant barcodes in hawthorn species will help to explain the genome structure of hawthorn.

Acknowledgements

This study was supported by TUBITAK (The Scientific and Technological Research Council of Türkiye) and Research Fund of the Ataturk University. Project Numbers are 121O217 (TUBITAK) and FHD-2022-11030 (Ataturk University)

Interest conflict

The authors declare no conflict of interest.

Author's contribution

H. I. S., S. E., M. A. and E. I. contributed to writing and editing the manuscript. R. A., A. G. K. and S. M. were responsible for analysis. All authors have read and agreed to the published version of the manuscript.

References

1. Nas, MN. Cultivation of hawthorn (*Crataegus* spp.): Opportuni-

ties and challenges. Paper presented at the 1st National Hawthorn Workshop, Malatya, 2012;Türkiye, 3-8.
2. Ercisli S, Esitken A, Turkkal C, Orhan E. The allelopathic effects of juglone and walnut leaf extracts on yield, growth, chemical and PNE composition of strawberry cv. Fern. Plant Soil Environ 2005;51:283-387. https://doi.org/10.1080/14620316.2005.11511888
3. Ikinci A, Bolat I, Ercisli S, Kodad O. Influence of rootstocks on growth, yield, fruit quality and leaf mineral element contents of pear cv. 'Santa Maria' in semi-arid conditions. Biol. Res 2014;47:71. https://doi.org/10.1186/0717-6287-47-71
4. Erturk Y, Ercisli S, Cakmakci R. Yield and growth response of strawberry to plant growth-promoting rhizobacteria inoculation. J. Plant Nutr 2012;35(6):817-826. https://doi.org/10.1080/01904167.2012.663437
5. Eyduran SP, Ercisli S, Akin M, Eyduran E. Genetic characterization of autochthonous grapevine cultivars from Eastern Turkey by simple sequence repeats (SSRs). Biotechnol. Biotechnol. Equip 2016;30(1):26-31. https://doi.org/10.1080/13102818.2015.1105726
6. Orhan E, Eyduran SP, Poljuha D, Akin M, Weber T, Ercisli S. Genetic diversity detection of seed-propagated walnut (*Juglans regia* L.) germplasm from Eastern Anatolia using SSR markers. Folia Hortic 2020;32(1):37-46. https://doi.org/10.2478/fhort-2020-0004
7. Donmez AA. The genus *Crataegus* L. (Rosaceae) with special reference to hybridisation and biodiversity in Turkey. Turk. J. Bot 2004;28(1-2):29-37. https://journals.tubitak.gov.tr/botany/vol28/iss1/4
8. Ercisli S. A short review of the fruit germplasm resources of Turkey. Genet. Resour. Crop Evol 2004;51(4):419-435. https://doi.org/10.1023/B:GRES.0000023458.60138.79
9. Kurt O, Celik N, Gore M, Kurt H. Threats to biodiversity biotrafficking in Turkey. Turkish JAF Sci. Tech 2019;7:46-51.
10. Ozhatay N, Kultur S, Gurdal B. Check-list of additional taxa to the supplement flora of Turkey VIII. J. Pharm. Istanbul Univ 2017;47(1):31-46. https://doi.org/10.26650/IstanbulJPharm.2022.1096223
11. Casas AM, Igartua E, Balaguer G, Moreno MA. Genetic diversity of Prunus rootstocks analyzed by RAPD markers. Euphytica 1999;110:139-149. https://doi.org/10.1023/A:1003745311408
12. Zamani Z, Sarkhosh A, Fatahi R, Ebadi A. Genetic relationships among pomegranate genotypes studied by fruit characteristics and RAPD markers. J. Hortic. Sci. Biotechnol 2007;82:11-18. https://doi.org/10.1080/14620316.2007.11512192
13. Duminil J, Di Michele M. Plant species delimitation: A comparison of morphological and molecular markers. Plant Biosyst 2009;143(3):528-542. https://doi.org/10.1080/11263500902722964
14. Bozhuyuk MR, Ercisli S, Orhan E, Koc A. Determination of the genetic diversity of walnut (*Juglans regia* L.) cultivar candidates from Northeastern Turkey using SSR markers. Mitt. Klosterneubg 2020;70(4):269-277.
15. Mcclintock B. The origin and behavior of mutable loci in maize. Proc. Natl. Acad. Sci. USA 1950;36(6):344-355. https://doi.org/10.1073/pnas.36.6.34
16. Papolu PK, Ramakrishnan M, Mullasseri S, Kalendar R, Wei Q, Ahmad Z, *et al*. Retrotransposons: How the continuous evolutionary front shapes plant genomes for response to heat stress. Front. Plant Sci 2022;13. https://doi.org/10.3389/fpls.2022.1064847
17. Beulé T, Agbessi MD, Dussert S, Jaligot E, Guyot R. Genome-wide analysis of LTR-retrotransposons in oil palm. BMC Genom 2015;16(1):1-14. https://doi.org/10.1186/s12864-015-2023-1.
18. Bennetzen JL, Ma J, Devos KM. Mechanisms of recent genome size variation in flowering plants. Ann. Bot 2005;95(1):127-132. https://doi.org/10.1093/aob/mci008
19. Boluk E. Determination, identification, reactions in some rice cultivars and determination of transposon gene movements of rice bacterial diseases. Master dissertation, Ondokuz Mayis University, Turkey, 2020.
20. Bourque G, Burns KH, Gehring M, Gorbunova V, Seluanov A, Hammell M, *et al*. Ten things you should know about transposable elements. Genome Biol 2018;19(1):1-12. https://doi.org/10.1186/s13059-018-1577-z
21. Kalendar R, Antonius K, Smýkal P, Schulman AH. iPBS: a universal method for DNA fingerprinting and retrotransposon isolation, Theor. Appl. Genet 2010;121(8):1419-1430. https://doi.org/10.1007/s00122-010-1398-2
22. Amiteye S. Basic concepts and methodologies of DNA marker systems in plant molecular breeding. Heliyon 2021;7(10):e08093. https://doi.org/10.1016/j.heliyon.2021.e08093
23. Collard BC, Mackill DJ. Start codon targeted (SCoT) polymorphism: a simple, novel DNA marker technique for generating gene-targeted markers in plants. Plant Mol. Biol. Rep, 2009;27(1):86-93. https://doi.org/10.1007/s11105-008-0060-5
24. Espinosa-Leal CA, Puente-Garza C A, García-Lara S. In vitro plant tissue culture: means for production of biological active compounds. Planta, 2018;248(1):1-18. https://doi.org/10.1007/s00425-018-2910-1
25. Buddhachat K, Thongphueak J, Changtor P, Nukool W, Saengthong M, Kamol P, *et al*. Herbal species authentication by melting fingerprint coupled with high resolution melting analysis (MF-HRM). Mol. Biol. Rep 2023;1-13. https://doi.org/10.1007/s11033-022-08235-x
26. Luo C, He XH, Chen H, Ou SJ, Gao MP, Brown JS, *et al.* Genetic diversity of mango cultivars estimated using SCoT and ISSR markers. Biochem Syst. Ecol 2011;39:676–684. https://doi.org/10.1016/j.bse.2011.05.023
27. Yan L, Ogutu C, Huang L, Wang X, Zhou H, Lv Y, *et al*. Genetic diversity and population structure of coffee germplasm collections in China revealed by ISSR markers. Plant Mol. Biol. Rep 2019;37(3):204-213. https://doi.org/10.1007/s11105-019-01148-3
28. Kurban M. Determination of the Saccharomyces cerevisiae dna barcode and creating a data base. PhD dissertation, Hacettepe University, 2019;Turkey, pp. 80.
29. Mishra, P., Kumar, A., Nagireddy, A., Mani, D. N., Shukla, A. K., Tiwari, R., *et al*. DNA barcoding: an efficient tool to overcome authentication challenges in the herbal market. *Plant Biotechnol. J* 2016;14(1):8-21. https://doi.org/10.1111/pbi.12419
30. Kress WJ, Erickson DL. A two-locus global DNA barcode for land plants: the coding rbcL gene complements the non-coding trnH-psbA spacer region. PLoS one 2007;2(6):e508. https://doi.org/10.1371/journal.pone.0000508
31. Ford CS, Ayres KL, Toomey N, Haider N, Van Alphen Stahl J, Kelly LJ, *et al*. Selection of candidate coding DNA barcoding regions for use on land plants. Bot. J. Linn. Soc 2009; 159(1):1-11. https://doi.org/10.1111/j.1095-8339.2008.00938.x
32. De Vere N, Rich TCG, Trinder SA, Long C. DNA barcoding for plants. In: Batley, J. (eds) plant genotyping. Methods Mol. Biol 2015;vol 1245. Humana Press, New York, NY. https://doi.org/10.1007/978-1-4939-1966-6_8
33. Emami A, Shabanian N, Rahmani MS, Khadivi A, Mohammad-Panah N. Genetic characterization of the *Crataegus* genus: implications for in situ conservation. Sci. Hortic 2018;231:56-65. https://doi.org/10.1016/j.scienta.2017.12.014
34. Cheng T, Xu C, Lei L, Li C, Zhang Y, Zhou S. Barcoding the kingdom Plantae: new PCR primers for ITS regions of plants with

improved universality and specificity. Mol. Ecol. Resour 2016; 16(1):138-149. https://doi.org/10.1111/1755-0998.12438

35. Shen WK, Xu GH, Luo MZ, Jiang ZD. Genetic diversity of *Sporisorium scitamineum* in mainland China assessed by SCoT analysis. Trop. Plant Pathol 2016;41(5):288-296. https://doi.org/10.1007/s40858-016-0099-z
36. Tyagi R, Sharma V, Sureja AK, Das Munshi A, Arya L, Saha D, *et al*. Genetic diversity and population structure detection in sponge gourd (*Luffa cylindrica*) using ISSR, SCoT and morphological markers. Physiol. Mol. Biol. Plants 2020;26(1):119-131. https://doi.org/10.1007/s12298-019-00723-y
37. Mansoory A, Khademi O, Naji AM, Rohollahi I, Sepahvand E. Evaluation of genetic diversity in three diospyros species, collected from different regions in Iran, using ISSR and SCoT molecular markers. Int. J. Fruit Sci 2022;22(1):235-248. https://doi.org/10.1080/15538362.2022.2034563
38. Mavlyutov Y, Kostenko S, Shamustakimova A, Klimenko I. Genetic variability analysis of Russian cultivars of ryegrass (*Lolium*) based on SCoT markers. J Genet Eng. Biotechnol 2022;20 (1):1-8. https://doi.org/10.1186/s43141-022-00446-w
39. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S. *et al*. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinform 2012;28(12):1647-1649. https://doi.org/10.1093/bioinformatics/bts199
40. Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Molecular biology and evolution, 2016;33(7):1870-1874. https://doi.org/10.1093/molbev/msw054
41. Mayer, Christoph, Phobos 3.3.11, 2006-2010. https://www.ruhr-uni-bochum.de/ecoevo/cm/cm_phobos.htm
42. Larkin MA, Blackshields G, Brown NP, Chenna R, Mcgettigan PA, Mcwilliam H. *et al*. Clustal W and Clustal X version 2.0. Bioinform 2007;23(21):2947-2948. https://doi.org/10.1093/bioinformatics/btm404
43. Dice LR. Measures of the amount of ecologic association between species. Ecol 1945;26(3):297-302. https://doi.org/10.2307/1932409
44. Kimura M, Crow JF. The number of alleles that can be maintained in a finite population. Genetics 1964; 49(4):725. https://doi.org/10.1093/genetics/49.4.725
45. Lewontin RC. Testing the theory of natural selection. Nature 1972;236 (5343), 181-182.
46. Yeh FC. POPGENE (version 1.3. 1). Microsoft Window-Bases Freeware for Population Genetic Analysis. *Cini Artic* 1999; https://ci.nii.ac.jp/naid/10025361630/
47. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics 2000;155(2):945-959. https://doi.org/10.1093/genetics/155.2.945
48. Peakall R, Smouse PE. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol. Ecol. Notes 2006;6(1):288-295. https://doi.org/10.1111/j.1471-8286.2005.01155.x
49. Nei M. Genetic distance between populations. Am. Nat 1972;106(949):283-292. https://doi.org/10.1086/282771
50. Feschotte C, Wessler SR. Mariner-like transposases are widespread and diverse in flowering plants. Proc. Natl. Acad. Sci 2002;99:280–285. https://doi.org/10.1073/pnas.022626699
51. Schulman AH, Flavell AJ, Ellis TN. The application of LTR retrotransposons as molecular markers in plants. Mob. Genet. Elem: Protoc. Genom. App 2004;145-173. https://doi.org/10.1385/1-59259-755-6:145
52. Gurlen A, Gundogdu M, Ozer G, Ercisli S, Duralija B. Primary, secondary metabolites and molecular characterization of hawthorn (*Crataegus* spp.) genotypes. Agronomy 2020;10(11):1731. https://doi.org/10.3390/agronomy10111731
53. Ferrazzini D, Monteleone I, Belletti, P. Small-scale genetic diversity in oneseed hawthorn (*Crataegus monogyna* Jacq.). Eur. J. For. Res 2008;127(5):407-414. https://doi.org/10.1007/s10342-008-0224-8
54. Erfani-Moghadam J, Mozafari M, Fazeli A. Genetic variation of some hawthorn species based on phenotypic characteristics and RAPD marker. Biotechnol. Biotechnol. Equip 2016;30(2):247-253. https://doi.org/10.1080/13102818.2015.1121790
55. Rahmani MS, Shabanian N, Khadivi-Khub A, Woeste KE, Badakhshan H, Alikhani L. Population structure and genotypic variation of *Crataegus pontica* inferred by molecular markers. Gene 2015; 572(1):123-129. https://doi.org/10.1016/j.gene.2015.07.001
56. Brown JA, Beatty GE, Finlay C, Montgomery WI, Tosh DG, Provan J. Genetic analyses reveal high levels of seed and pollen flow in hawthorn (*Crataegus monogyna* Jacq.), a key component of hedgerows. Tree Genet. Genom 2016;12(3):1-8. https://doi.org/10.1007/s11295-016-1020-0
57. Yildiz E, Sümbül A, Yaman M, Nadeem MA, Say A, Baloch FS, Popescu GC. Assessing the genetic diversity in hawthorn (*Crataegus* spp.) genotypes using morphological, phytochemical and molecular markers. Genet Resour Crop Evol 2023;70:135–146. https://doi.org/10.1007/s10722-022-01414-6.
58. Zhang X, Sun X, Wang J, Xue M, Sun C, Dong W. Evaluation of molecular and phenotypic diversity of *Crataegus bretschneideri* C. K. Schneid. and related species in China. Genet Resour Crop Evol 2023;70:221–234. https://doi.org/10.1007/s10722-022-01428-0.
59. Zarrei M, Talent N, Kuzmina M, Lee J, Lund J, Shipley PR, Dickinson TA. DNA barcodes from four loci provide poor resolution of taxonomic groups in the genus *Crataegus*. AoB Plants 2015;7. https://doi.org/10.1093/aobpla/plv045
60. Shin D, Kang HS, Park EM, Kim J, Kwon J, Suh J, *et al*. Authentication of tejocote (*Crataegus mexicana*) dietary supplements based on DNA barcoding and chemical profiling. Food Addit. Contam 2021;Part A, 38(12):1985-1994. https://doi.org/10.1080/19440049.2021.1964701.
61. Dogan H, Ercisli S, Temim E, Hadziabulic A, Tosun M, Yilmaz SO, Zia-Ul-Haq M. Diversity of chemical content and biological activity in flower buds of a wide number of wild grown caper (*Capparis ovate* Desf.) genotypes from Turkey. C. R. Acad. Bulg. Sci 2014, 67 (11): 1593-1600.
62. Abanoz YY, Okcu Z. Biochemical content of cherry laurel (*Prunus laurocerasus* L.) fruits with edible coatings based on caseinat, Semperfresh and lecithin. Turk. J. Agric. For 2022, 46 (6): 908-918. https://doi.org/10.55730/1300-011X.3052.
63. Duran ST, Ipek A. Identification of SNP markers linked to Rf locus in carrot using GBS. Turk. J. Agric. For 2022, 46 (6): 898-907. https://doi.org/10.55730/1300-011X.3051.
64. Yildiz M, Arbizu C. Inter-primer binding site (iPBS) retrotransposon markers provide insights into the genetic diversity and population structure of carrots (Daucus, Apiaceae). Turk. J. Agric. For 2022, 46 (2), 214-223. https://doi.org/10.55730/1300-011X.2972
65. Hajizadeh HS, Mortazavi SN, Ganjinajad, M, Okatan V, Kahramanoğlu İ. Evaluation of the optimum threshold of gamma-ray for inducing mutation on Polianthes tuberosa cv. Double and analysis of genetic variation with RAPD marker. Int J Rad Biol 2023, (just-accepted), 1-26. https://doi.org/10.1080/09553002.2023.2159566
66. Panfilova O, Kahramanoğlu I, Ondrasek G, Okatan V, Ryago

N, Tsoy M, Golyaeva O, Knyazev S. Creation and use of highly adaptive productive and technological red currant genotypes to improve the assortment and introduction into different ecological and geographical zones. Plants. 2022 Mar 17;11(6):802. https://doi.org/10.3390/plants11060802