



Research article

Assessment of genetic diversity on tomato (*Lycopersicon esculentum* Mill.) landraces using SSR molecular markers in Turkey

Ahmet Okumus^{*1} , Senay Dagidir² 

¹Adnan Menderes University, Agriculture Faculty, Agricultural Biotechnology Department, 09010, Aydin, Turkey

²Ondokuz Mayıs University, Science Faculty, Biology Department, 55100, Samsun, Turkey

Abstract

Genetic diversity is getting have been increasingly narrowed by the dense bred cultivar usage for production, on the other hand, genetic variation created by the landraces is known as the raw material of plant breeding. A collection of 21-landrace and three-commercial tomato (*Lycopersicon esculentum* Mill.) genotypes were screened for allele profile, genetic definition and genetic similarity with the aim of identifying genetic variability in genetic level with 40 SSR loci. Number of alleles per locus ranged from 3 (SSR50, SSR80, SSR9, LEat002) to 13 (LEat018). Expected and observed heterozygosity values were scored between He: 0.753 and Ho: 0.714 respectively and the mean value of polymorphism information content (PIC) value of these loci was found as 0.694. Considering distribution of alleles at loci, the highest allele frequency was observed from SSR75 with % 62.5 LEat014 with % 58.3 LEga004 with % 50.0 locus, while the lowest was in AI491065 with 16.6% LEta024 locus. UPGMA (Unweighted Pair-Group Method using Arithmetic Means) method was conducted for cluster analysis. The dendrogram was consisted 2 main groups; Group 2 was the largest and contained many sub-groups. The highest genetic similarity level of genotypes from Bafra-Merkez and Sivas-Yukarıkale Village was found as 0.776. The lowest similarity ratio was observed between Sivas-Yıldızeli-Emirler Village and Samsun-Carsamba genotypes with 0.136. The obtained results are indicated that genetic diversity information of landraces with effective SSR loci will help to manage tomato genetic resources for tomato breeding.

Keywords: Genetic diversity; landraces; PIC value; *Solanum lycopersicum* L.; SSR; tomato

1. Introduction

Tomato (*Lycopersicon esculentum* Mill. $2n = 2 \times = 24$), one of the important species of the *Solanaceae* family that first appeared in the region of Andes Mountains in South America, is one of the most cultivated vegetables in the world. It is known that today, the cultivated tomatoes are developed by using *L. hirsutum*, *L. peruvianum* and *L. pimpinellifolium* (Cox, 2000; Vural et al., 2000). Tomato, which has a 100-year history when it has been taken into culture, has gained a great popularity especially in the last 25 years for food industry. The arrival of tomato varieties to Europe was made by Spanish and Portuguese merchants in the 16th century through seeds (Osma et al., 2012;

Fentik, 2017). Probably the first tomato entered to Europe was yellow. Because it is named as golden apple (pomodoro) in Italy. In Europe, for a long time, tomato has not been given importance due to the belief that it is poisonous. This situation continued until the 18th century and it was used as food after the testimony of botanists and gourmets in England and France (Wien, 1987). Whereas, it was brought to Anatolia almost 150 years ago (in 1900's) and is now widely grown and consumed admirably (Yazgan and Fidan, 1996). Important tomato producers in the world are; China, India, Turkey and the European Union (EU) countries. Turkey is the world's largest tomato producing country after India and China, realizes 7.11% of the world production (Tridge, 2021). Also, it realizes 3.4% of world

* Corresponding author.

E-mail address: ahmet.okumus@adu.edu.tr (A. Okumus).

<https://doi.org/10.51753/flsrt.957055> Author contributions

Received 24 June 2021; Accepted 12 August 2021

Available online 30 August 2021

2718-062X © 2021 This is an open access article published by Dergipark under the [CC BY](https://creativecommons.org/licenses/by/4.0/) license.

exports and ranks 8th (World's Top Exports, 2021).

Increasing the yield of tomatoes per unit area in the world goes through the use of productive and disease-resistant varieties in addition to planting and maintenance. Plant genetic resources have played an important role in increasing yield and disease resistance. Breeding studies, in which new varieties are developed, gain importance with the adequate evaluation of plant genetic resources. Plant genetic resources; the village populations defined as local varieties consist of their wild relatives, unused old varieties and genotypes with clearly defined hereditary characteristics. In particular, protection of wild species is extremely important for future plant breeding studies. These valuable resources are in danger of decreasing or even disappearing due to environmental and other pressures in the regions where they are located. Culture varieties have become homogeneous in terms of gene structures and contain much less genetic diversity than primitive forms and their wild relatives. Wild species, on the other hand, are gene stores that have a wide genetic basis and constitute an important resource in solving the problems that may arise in the future of cultivated plants or in gaining new features to the plants (Ozgen et al., 1995; Can et al., 2019; Hocaoglu-Ozyigit et al., 2021).

However, in a study conducted in the Netherlands, it was stated that as a result of today's breeding studies, more diversity occurred compared to the 1950s (Schouten et al., 2019). Due to the increased awareness of the need for genetic diversity in tomatoes and the benefits of vegetables and fruits grown in local resources, more attention has recently been given to the restoration of genetic diversity in tomato species (Henareh et al., 2015; 2016; Castellana et al., 2020). Molecular techniques are the most used technologies in recent years to determine genetic diversity in breeding studies on tomato plants. Especially, with SSR markers genetic variation has been quite informative in order to detect diversity in tomato genome. It has been proved by various researchers that SSR technique can be used in the purity tests of tomato breeding lines in determining genetic relationships, characterizing and identifying varieties (Kaemmer et al., 1995; Tam et al., 2005; Garcia-Martinez et al., 2006).

The Solanacea Genome Project (SGN) is trying to reveal the tomato genome and approximately 40% of the tomato genome is complete (Sol Genomics, 2021). This study will be able to shed light on the future studies of the National Gene Bank by applying the SSR technique, which is used to create a gene bank in the world and collect varieties in a database He et al. (2003), developed and characterized SSR markers, 158 pairs of SSR primers were screened in 19 different tomato set varieties. They obtained 129 pairs of DNA fragments were and 65 of them were found to be polymorphic, and 19 sets of tomato varieties were classified according to their polymorphic SSR loci. Garcia-Martinez et al. (2006), used combinations of 19 SSR markers and 7 AFLP primers to characterize 48 local tomato varieties of south-eastern Spain and reported that SSR and AFLP markers were effective in the identification of local cultivated tomatoes. Tam et al. (2005) used SSAP, AFLP and SSR methods to determine genetic relatedness in commercially important tomatoes and peppers. In SSR method, they used 16 primer sets for tomato and 13 primer sets for pepper. They scored 39 different alleles for tomato and 31 different alleles for pepper. They found that the total number of alleles per primary in tomato as 2.44 and the number of polymorphic alleles as 2.44 (100%). In pepper, the number of alleles per primary by SSR method was 2.385 and the number of polymorphic alleles was 2.385 (100%). They reported that the SSR method is a method reveals the

specific characteristics of the plant. In order to determine genetic diversity in 39 inbred tomato lines, Benor et al. (2008) used 35 types of SSR polymorphic markers collected from China, Japan, South Korea and USA. They determined that these tomato lines, from which 150 alleles were obtained, showed moderate variation, and some varieties gave unique alleles. They reported that the number of alleles per locus was 4.3 and the average polymorphism information content (PIC) was 0.31. They also stated that there is a separate and more distant kinship in the USA variety, which is one of the 39 tomato lines.

Although Turkey it is not included in the center or origin of the Solanaceae family, which tomato is included, has been obtaining seed production by farmers. As a result, it has adapted to different conditions and a natural gene pool containing different types has been formed (Dilbirligi, 2007). In this research, in order to reveal the genetic diversity of the different regions of Turkey, the seeds produced by farmers were selected. It was aimed to make identification at the DNA level, to reveal the genetic relationships (similarities or differences) and to determine some morphological characteristics using 40 SSR primers on a total of 24 tomato varieties.

2. Materials and methods

2.1. Plant materials

This research was carried out in the greenhouse and Agricultural Biotechnology Laboratory of Ondokuz Mayıs University Faculty of Agriculture, Department of Agricultural Biotechnology. In the study, 21 local tomato varieties collected from different regions of our country, which are thought to be different from each other, and three commercial tomato varieties were used (Table 1).

Table 1

The genotype codes and their collected regions.

No	Genotype Codes	Collected Regions
1	GRC	GIRESUN-CAVUSLU
2	UNS	UNYE-SOFUTEPESEI
3	UNM	UNYE
4	SMM1	SAMSUN
5	SVYE	SIVAS-YILDIZELI-EMIRLER
6	SVY1	SIVAS-YUKARIKALE
7	SVK	SIVAS-KOYULHISAR
8	SVGK	SIVAS-GEMEREK-KUMEOREN
9	BFM1	BAFRA
10	NDM	NIGDE
11	SVA	SIVAS-AKKUZULU
12	SVY2	SIVAS-YUKARIKALE
13	SVY3	SIVAS-YUKARIKALE
14	KRSB	KARABUK-SAFRANBOLU-BOSTANBUKU
15	AMGS	AMASYA-GUMUSHACIKOY-SALLAR
16	ANKC	ANTALYA-KAS-CAVDIR
17	SME	SAMSUN-ENGIZ
18	SMC	SAMSUN-CARSAMBA
19	MLBM	MUGLA-BODRUM-MUMCULAR
20	SMCI	SAMSUN-CINARLIK
21	AYNP	AYDIN-NAZILLI-PIRLIBEY
22	UCW1	COMMERCIAL
23	UCW2	COMMERCIAL
24	UCW3	COMMERCIAL

While the genotype codes used in the study were prepared, two consonants of the province from which the sample was taken, and the following ones define the district and village (for

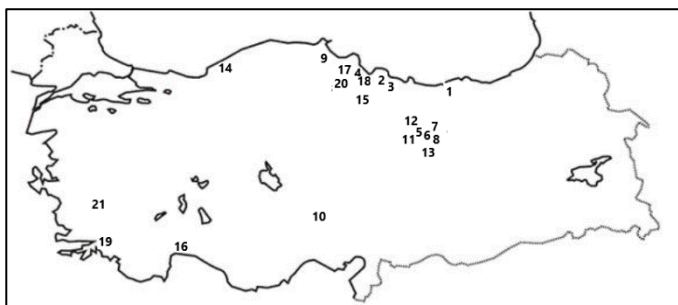


Fig. 1. The localities, where tomato genotypes are collected.

example; GRC: Giresun-Cavuslu), while commercial varieties were coded as UCW1, UCW2, and UCW3. The code numbers of the tomato genotypes used in the experiment and the places where they were collected are given in Fig. 1 and Table 1.

2.2. DNA isolation and PCR

DNA isolation of 24 tomato varieties studied in the study was performed according to the method of CTAB DNA isolation protocol developed by Doyle and Doyle (1991), while DNA quality and quantity measurements were made using 0.8% agarose gel and spectrophotometer (Biophotometer, Eppendorf, Hamburg, Germany). Thermal Cycle Device (PEQLAB Primus 96 Gradient PCR Thermal Cycler) was used for DNA reproduction and PCR optimization studies. In PCR amplification; a total of 25 µl PCR mix was prepared: 4 µl DNA, 2 µl primer 1, 2 µl primer 2, 8 µl TMastermix, 9 µl water. For DNA amplification, in PCR at 94 °C for 5 min, at 94 °C; 35 cycles 55 °C 1 min, 72 °C 2min, 94 °C 1 min last cycle was 55 °C 1 min and 72 °C 7 min protocol was applied.

2.3. Primers

It was made with primers that were selected from primer sets that showed the most allelic variation previously used in tomato genotypes. The selection of primers was made according to types that show a selective difference between populations with at least 3 alleles, and care was taken to consist of primers that can be used as a standard in seed purity and genetic differences, which will guide further research. Care was taken to distribute the primers as evenly as possible into the genome. The primer set used is given in Table 2. For this purpose, 40 SSR primer sets in tomato genome bank (Sol Genomics, 2021) were used.

2.4. Statistical analysis

The scoring process of gels, the data of the markers were coded in a binary system according to the presence and absence of the bands indicating alleles.

According to this, genetic parameters; number of alleles per locus (n), allele frequency, expected heterozygosity (He), observed heterozygosity (Ho), effective allele number (Ae) and polymorphism information content (PIC) of microsatellite markers used were also calculated. Dendrogram of genotypes was created and visualized with software program NTSYS (version 2.02g, Exeter Software, Setauket, NY). UPGMA (Unweighted Pair-Group Method using Arithmetic means) method was used for dendrogram. Statistical methods and formulas used to evaluate the data obtained from microsatellite markers that summarized below.

2.4.1. Expected heterozygosity (He)

The most commonly used measure of genetic variation in a population is heterozygosity. For haploid data, the $h = 1 - \sum p_i^2$ formula is used to calculate genetic diversity, while in codominant molecular marker studies, the formula $He = 1 - \sum p_i^2$ is used to calculate the expected heterozygosity (Peakall and Smouse, 2010). The pi value here shows the frequency of the “i” allele in the sample studied (Nei, 1987).

2.4.2. Observed heterozygosity (Ho)

The observed heterozygous value calculated per locus is the ratio between the heterozygous genotypes and the total genotype analyzed.

$Ho = \text{number of heterozygotes at a locus} / \text{total number of genotypes}$

2.4.3. Allele count

Another indicator of genetic variation is the number of alleles per locus (A). This criterion, also known as allelic richness, is affected by the number of samples (Nei, 1987).

$$\text{Average } Na = \sum nai / r$$

nai: number of alleles of locus i, r: number of loci

2.4.4. Effective number of alleles (Ne)

It is a strategic parameter used in measuring diversity in codominant molecular marker studies. This calculation provides meaningful comparisons of allelic diversity between different allele frequency distributions and locus, thus providing an estimate of the ideal allele number (Peakall and Smouse, 2010).

This criterion, developed by Kimura and Crow (1978), is the reciprocal of homozygousness.

$$Ne = 1 / \sum pi^2$$

Ne: effective number of alleles

xi: is the frequency of the i allele.

2.4.5. Polymorphism information content (PIC)

Polymorphism information content is a value that measures the usefulness and informative level of a marker and shows a structure dependent on the number of alleles determined and the frequency distribution of these alleles (Botstein et al., 1980). This value is calculated as $PIC = 1 - \sum pi^2 - \sum pi^2 pj^2$. Where pi and pj are the “i” and “j” alleles, respectively. All these statistical values are shown in Table 3.

2.4.6. Genetic similarity

Genetic distance is the magnitude of gene differences between species (or population) pairs. These values are generally equivalent to geometric distances, so a distance value of “0” indicates no difference.

Similarity (I) and distance (D) values are complementary to each other ($I + D = 1$).

Table 2

The primer sets used in this study.

NO	PRIMERS	REPEATS	SEQUENCES
1	SSR47	(at)14	F: tcc tca aga aat gaa gct ctg a R: cct tgg aga taa caa cca caa
2	LEtc002	(ac)3(tc)6 imp	F: ttc tca cac ctg cac aca cc R: agc ggg atg att aca gaa atg
3	SSR139	(aga)2, (gaa)7	F: tgg gta tgg gat tta cac caa R: aaa cga agg caa caa cga ag
4	Tom236-237	(at)16	F: gtt ttt tca aca tca aag agc t R: gga tag gtt tgg tta gtg aac t
5	SSR572	(tc)11	F: aat tca cct ttc ttc cgt cg R: tgc aaa gaa caa aga cgg tg
6	LEat014	(at)9	F: tgt gtt gcg tca tta cca cta aac R: ccc aac cac caa tac ttt cc
7	SSR50	(tc)6, (ccctc)2	F: ccg tga ccc tct tta caa gc R: ttg ctt tct tct tgg cca tt
8	U81996	(ta)14	F: agg ttg atg aaa gct aaa tct ggc R: caa cca cca atg ttc att aca aga c
9	TMS37	(ga)21(ta)20	F: cct tgc agt tga ggt gaa tt R: tca agc acc tac aat caa tca
10	LEta019	(ta)20	F: tgt aga taa ctt cct agc gac aat c R: acg gac gga tgg aca aat g
11	LEat018	(at)29 imp	F: cgg cgt att caa act ctt gg R: gcg gac ctt tgt ttt ggt aa
12	TMS26	(ga)20	F: ttc ggt tta ttc tgc caa cc R: gcc tgt agg att ttc gcc ta
13	AQ368062	(ta)19	F: tga tcc taa gct ttt tcc gtg agt R: caa gtt cac ctc att tca ccc ct
14	Y08306	(ta)11	F: aac ggt gga aac tat tga aag g R: cac cac caa acc cat cgt c
15	SSR80	(tttcaa)2, (gtacaa)2, (caa)7	F: ggc aaa tgt caa agg att gg R: agg gtc atg ttc ttg att gtc a
16	SSR9	(ata)10	F: ccc ttt gca agt tct tct tca R: ttc atg agc caa cat agg agg
17	SSR75	(aat)9	F: cca tct att atc ttc tct cca aca c R: ggt ccc aac tgg gta cac ac
18	AI773078	(aat)14	F: gat gga cac cct tca att tat ggt R: tcc aag tat cag gca cac cag c
19	AI491065	(at)9	F: act gca ttt cag gta cat act ctc R: ata aac tgg tag acc ata ccc tc
20	Y09371	(at)12	F: tga gaa caa cgt tta gag gag ctg R: cgg gca gaa tct cga act c
21	AI895126	(ta)9	F: gct ctg tcc tta caa atg ata cct cc R: caa tgc tgg gac aga aga ttt aat g
22	LEta017	(ta)5	F: gag cac cca tta att tgg tta cg R: gtg gcg gat cta gaa att taa act g
23	SSR136	(cag)7	F: gaa acc gcc tct ttc act tg R: cag caa tga ttc cag cga ta
24	LEat016	(at)9	F: ccc aaa tgc tat gca ata cac R: agt tca gga ttg gtt taa ggg
25	LEgt001	(at)17(gt)18	F: aga att ttt tca tga aat tgt cc R: tat tgc gtt cca ctc cct ct
26	LEta014	(ta)31(gata)13 imp	F: aca aac tca aga taa gta aga gc R: gtg aat tgt gtt tta aca tgg
27	LEta024	(tg)4(ta)5	F: taa ata caa aag cag gag tgg R: gag ttg aca gat cct tca atg
28	LEtat003	(gt)2(ta)3(tat)6 imp	F: cat ttt atc att tat ttg tgt ctt g R: aca aaa aaa ggt gac gat aca
29	LEtat002	(tat)12	F: acg ctt ggc tgc ctc gga R: aac ttt att att gcc acg tag tca tga
30	LEta023	(ga)24(ta)31 imp	F: att gct cat aca taa ccc cc R: ggg aca aaa tgg taa tcc at
31	LEta021	(ta)11	F: ttc ttc cgt atg agt gag t R: ctc tat tac tta tta tta tgg
32	LEta017	(ta)5	F: gag cac cca tta att tgg tta cg R: gtg gcg gat cta gaa att taa act g
33	LEta016	(ta)14	F: agg ttg atg aaa gct aaa tct ggc R: caa cca cca atg ttc att aca aga c
34	LEta012	(ta)19	F: tga tcc taa gct ttt tcc gtg agt R: caa gtt cac ctc att tca ccc ct
35	LEta007	(ta)20	F: gcc gtt ctt ggt gga tta g R: cct cct ttc gtc tct ttg tc
36	LEgata002	(gata)26	F: ttg gta att tat gtt cgg ga R: ttg agc caa ttg att aat aag tt
37	LEgata001	(gata)45	F: ctc tct caa tgt ttg tct ttc R: gca agg tag gta gct agg ga
38	LEga004	(ga)26 imp	F: agc atg gga aga aga cac gt R: ttg agc aaa aca tgg caa tc
39	AI780156	(ct)12	F: tcc aat ttc agt aag gac ccc tc R: ccg aaa acc ttt gct aca gag tag a
40	LEct004	(ct)3c14(ct)23	F: agc cac cca tca caa aga tt R: gtc gca cta ttc gtc acg ta

Genetic distance is a statistical measure that has been standardized as an evaluation scale in the investigation of genetic differences and provides an evaluation scale. The most commonly used genetic distance value is Nei's Genetic Distance (Nei, 1972). Genetic similarity between the two genotypes (I);

$$I = 2N_{XY} / N_X + N_Y$$

It is calculated by the formula (Nei and Li, 1979). Where N_{xy} ; the bands shared by the two genotypes (x and y); N_x and N_y show the individual band numbers of x and y. Nei and Li (1979) reported that genetic distance is the negative (-) natural logarithm of the measure of genetic similarity.

$$D = - \ln (I)$$

D; It means genetic distance.

In terms of 40 microsatellite markers used, the genetic similarity and distance conditions of the 24 studied genotypes were calculated using the NTSYSpc v2.11 program according to the above formulas (Rohlf, 1998).

3. Results and discussion

3.1. Genotypic parameters

Imaging and scoring of the gels was done on the Syngene-Gene Tools (Cambridge, UK) gel documentation device and the molecular sizes of the PCR products in bp were extracted. In the study, for the analysis of 24 tomato genotypes with 40 polymorphic SSR loci, the number of alleles at each locus (n), allele frequency, expected (H_e) and observed heterozygosity (H_o), predicted silent allele (null) frequency (r) and polymorphism information content (PIC) were determined (Table 3). Genetic difference was measured by calculating expected heterozygosity as $1 - \pi^2$. The π value here shows the frequency of the "i" allele in the sample studied (Nei, 1987). The observed heterozygosity is the ratio between the heterozygous genotypes and the total genotype analyzed. The estimation of the presence of the null allele was calculated as $(H_e - H_o) / (1 + H_e)$ (Brookfield, 1996). All these statistical values and parameters are shown in Table 3. Accordingly, the number of alleles, one of the components of genetic diversity, varied from 3 to 13, giving an average value of 6.4. The number of effective alleles (N_e)

Table 3
Statistical values and parameters of polymorphism information content.

No	Loci	Na	He	Ho	r	Ne	PIC
1	SSR47	4	0.698	0.584	0.067	3.3	0.600
2	LEttc002	7	0.723	0.719	0.002	3.6	0.718
3	SSR139	8	0.820	0.817	0.001	5.5	0.802
4	Tom236-237	5	0.745	0.671	0.042	3.9	0.640
5	SSR572	5	0.749	0.710	0.022	3.9	0.686
6	LEat014	4	0.546	0.625	-0.051	2.0	0.545
7	SSR50	3	0.605	0.595	0.068	2.5	0.455
8	U81996	10	0.867	0.859	0.004	7.5	0.866
9	TMS37	4	0.691	0.680	0.006	3.2	0.614
10	LEta019	7	0.788	0.666	0.068	4.7	0.708
11	LEat018	13	0.911	0.903	0.004	11.2	0.903
12	TMS26	6	0.747	0.750	-0.001	3.9	0.659
13	AQ368062	6	0.757	0.875	-0.067	4.7	0.745
14	Y08306	5	0.773	0.709	0.036	4.4	0.707
15	SSR80	3	0.653	0.417	0.142	2.8	0.540
16	SSR9	3	0.652	0.333	0.193	2.6	0.539
17	SSR75	4	0.558	0.702	-0.092	2.2	0.340
18	AI773078	4	0.648	0.789	-0.085	2.8	0.512
19	AI491065	10	0.880	0.875	0.002	8.3	0.879
20	Y09371	8	0.846	0.835	0.005	6.5	0.833
21	AI895126	4	0.670	0.626	0.026	3.0	0.554
22	SSR136	10	0.866	0.790	0.040	7.5	0.829
23	LEta017	8	0.852	0.750	0.055	6.7	0.809
24	LEat016	4	0.718	0.625	0.054	3.0	0.674
25	LEgt001	7	0.785	0.682	0.057	4.6	0.745
26	LEta014	9	0.830	0.792	0.020	5.8	0.823
27	LEta024	11	0.880	0.877	0.001	8.3	0.878
28	LEtat003	8	0.809	0.792	0.009	5.2	0.807
29	LEtat002	3	0.600	0.458	0.088	2.5	0.474
30	LEta023	9	0.818	0.824	-0.003	5.5	0.817
31	LEta021	6	0.804	0.755	0.027	5.1	0.735
32	LEata004	4	0.669	0.625	0.026	3.0	0.529
33	LEta012	4	0.679	0.628	0.030	3.1	0.561
34	LEta007	10	0.871	0.844	0.014	7.7	0.863
35	LEgata002	9	0.850	0.837	0.007	6.6	0.848
36	LEgata001	4	0.712	0.625	0.050	3.4	0.651
37	LEga004	5	0.642	0.638	0.002	2.8	0.645
38	AI780156	5	0.732	0.705	0.015	3.7	0.651
39	LEct004	11	0.878	0.874	0.002	8.2	0.870
40	Umc1178	6	0.798	0.709	0.049	4.9	0.736
	Loci	256	-	-	-	-	-
	Mean	6.4	0.753	0.714	-	4.752	0.694

Na: Number of alleles observed in the locus, **He:** Expected heterozygosity rate, **Ho:** Observed heterozygosity rate, **r:** Estimated null allele frequency, **Ne:** Effective allele number, **PIC:** Polymorphism information content.

varied from 11.2 to 2, with an average value of 4.75. While LEat018 locus gave the highest number of alleles, the lowest number of alleles were seen in SSR50, SSR80, SSR9 and LETat002 loci. In addition, these loci were found to have the highest and lowest effective allele values. 11.2 for the LEat018 marker; it was found as 2 for the LEat014 marker. These loci were also found to be the marker loci with the highest and lowest polymorphism information content values. The PIC value for the LEat018 marker was 0.903; PIC values for SSR50, SSR80, SSR9 and LETat002 markers were calculated as 0.455-0.540-0.539-0.474, respectively. The average PIC value of the loci used was found to be as 0.694.

The average expected and observed heterozygosity rates were found to be as 0.753 and 0.714, respectively. The predicted null allele frequency was negative at 6 loci and positive at 34 loci. In a study, PIC value at 37 loci was higher than the value of 0.05 (Sefc et al., 2001). This situation indicates that these selected microsatellite markers are indeed highly polymorphic

in tomato.

Looking at the loci individually, 24 genotypes gave LEat018 loci with 13 alleles at most, and this was LEta024, LEct004, (11 alleles), U81996, AI491065, SSR136, LEta007 (10 alleles), LEta014, LEta023, LEgata002, (9 alleles), SSR139, Y09371, LEta017, LETat003 (8 alleles), LEttc002, LEta019, LEgt001 (7 alleles), TMS26, AQ368062, LEta021, Umc1178 (6 alleles), Tom236-237, SSR572, Y08306, LEga004, AI 580) SSR47, LEat014, TMS37, SSR75, AI773078, AI895126, LEata004, LEat016, LEgata001, LEta012 (4 alleles) followed loci. The least alleles gave SSR50, SSR80, SSR9 and LETat002 as 3 alleles.

Allele numbers and frequencies of 40 loci examined in 24 genotypes are given in Table 3. Accordingly, at the LEat018 locus, 184, 186, and 195 alleles were the most common among genotypes giving 37.5% the highest frequency. In the study, when the loci used for the 24 genotypes were examined one by one, it was determined that the locus with the lowest number of

effective alleles (LEat014) was not among the loci with the lowest allele. Likewise, the most common allele in 24 genotypes were 186 (20.8%) at the LEta017 locus, 284.281 (29.2%) at the SSR572 locus, 237.235 (33.3%) at the TMS26 locus, 274.280 (29.1%) at the LEta019 locus, 288 (33.3%) at the LEgt001 locus, 206.214 (16.6%) at AI491065 locus, 196 (20.8%) at U81996 locus, 338 (33.3%) at LEgata001 locus, 201 (41.6%) at SSR80 locus, 198 (32.2%) at SSR139 locus, At LEata004 locus 184 (45.8%), at LEga004 locus 270 (50.0%), at LEta023 and LEta014 loci 178 (30.4% -29.16%), at LETtc002 locus 157 (45.8%), at LEta024 locus 290-288 (16.6%), at LETat003 locus 102 (29.1%), at Umc1178 locus 183-185 (25.0%), at AQ368062 locus 284 (37.5%), at locus Y08306 198 (29.2%), Tom236-195 (37.5%) at 237 locus, 172 (20.8%) at SSR136, 250 (29.1) at LEta021 locus, 364 (20.8%) at LEct004 locus, 184 (37.5%) at LEat016 locus, LEat014 at the locus 210 (58.3%), at the AI895126 locus 116 (41.6%), at the SSR50 locus 200 (50.0%), at the TMS37 locus 154 (33.3%), at the SSR75 locus 163 (62.5%), 264 (41.6%) at LEta012 locus, 187 (41.7%) at SSR9 locus, 160 (45.8%) at AI773078 locus, 314 (37.6%) at SSR47 locus, 295 (45.6%) at LETat002 locus, There were 312 (20.8%) alleles at the LEta007 locus, 354-349 (20.8%) at the LEgata locus, 175 (25.0%) at the Y09371 locus, and 118-123 (33.3%) at the AI780156 locus. Among the genotypes studied, SSR572 (284.281), TMS26 (237.235), LEta019 (274.280), LEgata001 (338), SSR80 (201), Tom 236-237 (195), LEat016 (184), SSR50 (200), SSR75 (163), Alleles with high rates at LEta012 (264), SSR9 (187), AI773078 (160), LEta007 (312), AI780156 (118-123) and other loci were also seen in commercial genotype varieties, including native genotype varieties. Therefore, we cannot use the expression that these are alleles belonging only to native genotype varieties. However, the most common 184-186 in LEat018 locus are at LEta017 locus, 288 at LEgt001 locus, 206-214 at AI491065 locus, 196 at U81996 locus, 198 at SSR139 locus, 184 at LEata004 locus, 270 at LEga004 locus, 178 at LEta023 and LEta014 locus, LETtc002 locus), 290-288 at LEta024 locus, 102 at LETat003 locus, 183-185 at Umc1178 locus, 284 at AQ368062 loci, 198 at Y08306, 172 at SSR136, 250 at LEta021 locus, 364 at LEct004 locus, 210 at LEat014 locus 154, 116 at SSR37 locus 314 at locus, 295 at LETat002 locus, 175 at Y09371 locus were only seen in native genotype varieties.

The similarity ratio index was found by calculating Genetic Distance $D = 1 - (\text{shared allele ratio})$ (uniqueness ratio, dissimilarity) with NTSYSpc v2.11 (Rohlf, 2000) program. This value was then transformed into similarity ratio. Genetic similarity values of genotypes are given in Table 4.

Accordingly, the highest genetic similarity was found between SVY2 [12] and BFM1 [9] (0.776), SMÇ [18] and AYNP [21] (0.698) and CRSB [14] and ANKÇ [16] (0.604). The shared allele ratio ranged from SVYE [5] to SME [17] with a value of at least 0.211. Genetic similarity value of commercial genotypes were found as UCW1 [22] and UCW2 [23] (0.578), UCW1 [22] and UCW3 [24] (0.498), UCW2 [23] and UCW3 [24] (0.317). Therefore, the result is that these varieties are different.

The greatest genetic similarity with commercial genotypes was found to be UCW1 [22] and SVGK [8] (0.589), UCW2 [23] and GRÇ [1] (0.568) and UCW3 [24] and SVY2 [12] (0.544). The lowest genetic similarity with commercial genotypes was found as UCW3 [24] and KRSB [14] (0.362), UCW2 [23] and BFM1 [14] (0.241) and UCW1 [22] and SMÇ1 [20] (0.217). Dendrogram NTsys of genotypes (version 2.02g. Exeter Soft-

Table 4
Genetic similarity values of tomato genotypes.

1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
1	1	.317	.423	.287	1	.468	.392	.574	1	.498	.438	.408	.423	.377	.377	1	.468	.392	.408	.423	.362	.272	.256	.256	.256
2	.423	1	.468	.392	.574	1	.498	.438	.408	.423	.377	.377	1	.468	.392	.408	.423	.362	.272	.256	.256	.256	.256	.256	.256
3	.287	.468	1	.498	.438	.408	.423	.377	.377	1	.468	.392	.408	.423	.362	.272	.256	.256	.256	.256	.256	.256	.256	.256	.256
4	.468	.392	.574	1	.498	.438	.408	.423	.377	.377	1	.468	.392	.408	.423	.362	.272	.256	.256	.256	.256	.256	.256	.256	.256
5	.498	.438	.408	.423	1	.468	.392	.408	.423	.377	.377	1	.468	.392	.408	.423	.362	.272	.256	.256	.256	.256	.256	.256	.256
6	.438	.453	.423	.377	.377	1	.468	.392	.408	.423	.362	.272	.256	.256	.256	.256	.256	.256	.256	.256	.256	.256	.256	.256	.256
7	.468	.392	.408	.423	.362	.272	1	.468	.392	.408	.423	.362	.272	.256	.256	.256	.256	.256	.256	.256	.256	.256	.256	.256	.256
8	.483	.438	.513	.513	.468	.453	.287	1	.468	.392	.408	.423	.362	.272	.256	.256	.256	.256	.256	.256	.256	.256	.256	.256	.256
9	.453	.528	.498	.513	.453	.287	.302	.513	1	.468	.392	.408	.423	.362	.272	.256	.256	.256	.256	.256	.256	.256	.256	.256	.256
10	.468	.513	.498	.498	.453	.272	.332	.241	.166	1	.468	.392	.408	.423	.362	.272	.256	.256	.256	.256	.256	.256	.256	.256	.256
11	.513	.528	.498	.513	.468	.392	.408	.317	.256	.196	1	.468	.392	.408	.423	.362	.272	.256	.256	.256	.256	.256	.256	.256	.256
12	.483	.513	.559	.528	.528	.362	.377	.317	.776	.241	.256	1	.468	.392	.408	.423	.362	.272	.256	.256	.256	.256	.256	.256	.256
13	.498	.513	.498	.568	.468	.408	.423	.362	.392	.347	.408	.302	1	.468	.392	.408	.423	.362	.272	.256	.256	.256	.256	.256	.256
14	.528	.498	.513	.544	.438	.528	.468	.498	.498	.498	.408	.438	.438	1	.468	.392	.408	.423	.362	.272	.256	.256	.256	.256	.256
15	.559	.528	.528	.559	.377	.574	.483	.498	.498	.498	.408	.438	.438	.362	.332	1	.468	.392	.408	.423	.362	.272	.256	.256	.256
16	.528	.574	.574	.423	.513	.559	.544	.544	.559	.544	.483	.544	.513	.604	.347	.347	1	.468	.392	.408	.423	.362	.272	.256	.256
17	.513	.544	.498	.544	.211	.528	.513	.559	.544	.513	.498	.513	.498	.468	.468	.332	.332	1	.468	.392	.408	.423	.362	.272	.256
18	.528	.513	.544	.574	.544	.559	.528	.528	.574	.559	.544	.513	.528	.513	.423	.392	.256	.256	1	.468	.392	.408	.423	.362	.272
19	.574	.528	.528	.574	.544	.589	.559	.559	.574	.574	.528	.528	.528	.528	.513	.438	.332	.136	.136	1	.468	.392	.408	.423	.362
20	.559	.544	.528	.544	.544	.528	.559	.528	.574	.574	.544	.528	.528	.483	.468	.453	.423	.317	.256	.256	1	.468	.392	.408	.423
21	.574	.559	.513	.574	.513	.589	.528	.544	.544	.559	.528	.513	.483	.453	.468	.468	.302	.226	.226	.226	.226	1	.468	.392	.408
22	.528	.544	.528	.438	.528	.513	.544	.589	.544	.544	.559	.559	.544	.559	.468	.483	.483	.483	.498	.498	.498	.498	1	.468	.392
23	.568	.544	.544	.559	.513	.498	.528	.574	.241	.528	.544	.544	.528	.513	.528	.483	.528	.513	.528	.513	.528	.498	.498	.498	.498
24	.272	.438	.513	.513	.528	.453	.528	.528	.483	.498	.513	.544	.483	.362	.408	.513	.528	.513	.528	.513	.483	.408	.498	.423	.423

ware. Setauket. NY created and visualized with software program. UPGMA (Unweighted Pair-Group Method using arithmetic means) method was applied for cluster analysis.

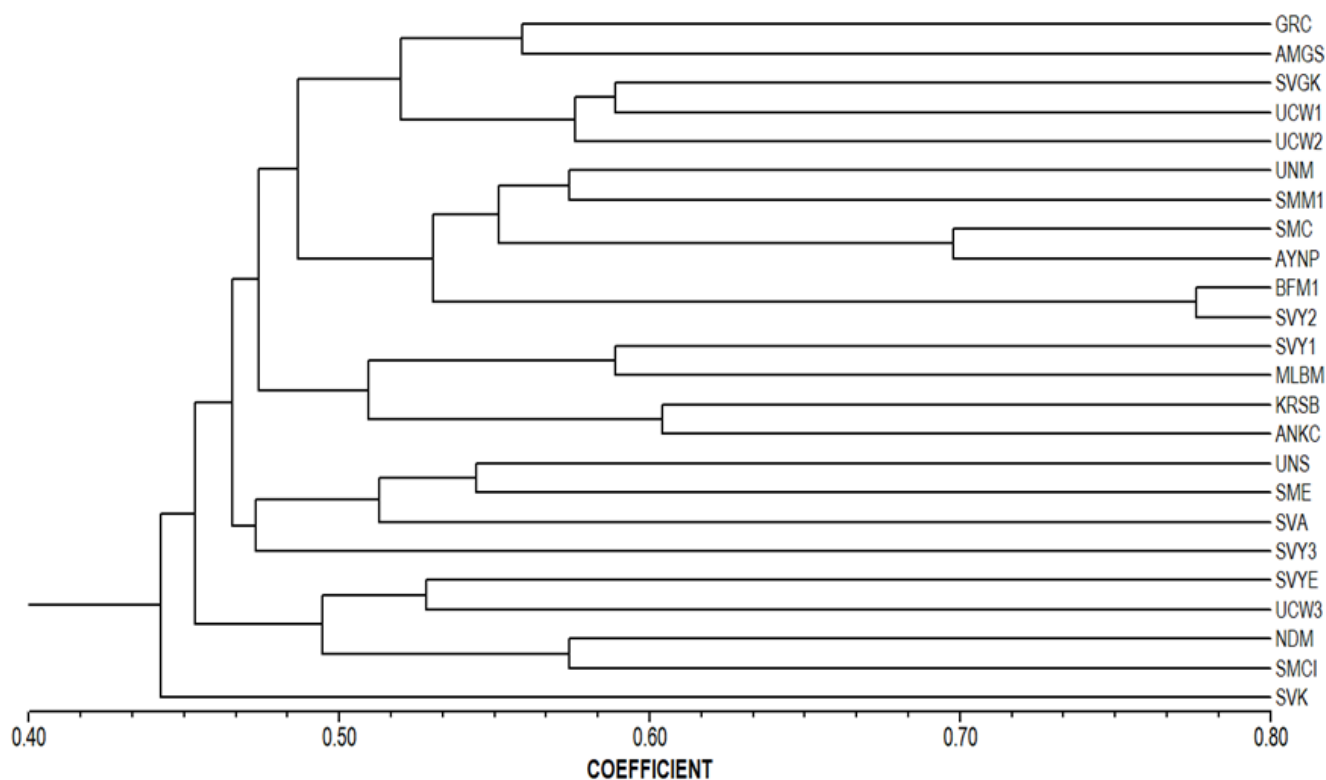


Fig. 2. Genetic diversity dendrogram between tomato genotypes.

Table 5

Morphological characteristics of the tomato genotypes.

Genotype Codes	Plant Type	Fruit Width-Length (mm)	Locule Numbers of Fruit	Maturity Time	Fruit Type
1-GRC	Determinate	12.39-12.92	2	Early	Round
2-ÜNS	Indeterminate	57.02-52.89	3	Medium	Round
3-ÜNM	Determinate	65.74-52.70	6	Late	Slightly Flattened
4-SMM1	Indeterminate	94.43-56.81	11	Late	Flattened
5-SVYE	Indeterminate	81.48-55.50	8	Late	Flattened
6-SVY1	Indeterminate	12.43-19.47	2	Late	Cylindrical
7-SVK	Determinate	65.45-62.79	8	Late	Pear-shaped
8-SVGK	Determinate	25.35-27.09	2	Medium	Round
9-BFM1	Determinate	97.79-60.13	8	Early	Flattened
10-NDM	Determinate	43.70-62.76	2	Early	Oval
11-SVA	Indeterminate	92.13-56.02	11	Medium	Flattened
12-SVY2	Determinate	84.12-72.22	9	Medium	Flattened
13-SVY3	Determinate	11.72-16.23	2	Medium	Oval
14-KRSB	Determinate	51.81-73.90	2	Medium	Oval
15-AMGS	Determinate	92.13-57.26	13	Medium	Flattened
16- ANKÇ	Determinate	49.50-3.83	4	Medium	Oval
17- SME	Determinate	49.81-43.29	2	Medium	Slightly Flattened
18-SMÇ	Determinate	52.60-49.72	3	Medium	Round
19-MLBM	Determinate	51.65-48.82	3	Early	Round
20-SMÇI	Determinate	85.30-70.55	6	Late	Round
21-AYNP	Determinate	72.77-53.26	5	Medium	Slightly Flattened
22-UCW1	Indeterminate	45.35-37.12	2	Early	Slightly Flattened
23-UCW2	Indeterminate	41.54-38.72	2	Early	Round
24-UCW3	Determinate	94.72-57.08	10	Medium	Flattened

Dendrogram of tomato genotypes are given in Fig. 2. When the dendrogram is examined, 2 main groups stand out. There were 5 subgroups within the second (2) group, including the majority of genotypes, and 3 subgroups within the first (1) group. Commercial (22), Commercial (23), Commercial (24) varieties used as commercial varieties have been observed to

form a separate subgroup within the first (1) group. In the dendrogram, the genetic similarities (0.776) of the Bafra-Merkez and Sivas-Yukarıkale Village cultivars assessed by the shared allele ratios are clearly seen. Again, a close genetic similarity value of 0.604 between Samsun-Carsamba and Aydin-Nazilli (0.698), which are genetically close, and Karabuk-

Safranbolu and Antalya-Kas, is also seen in the dendrogram. Sivas-Yildizeli, which is linked to Unye-Merkez and Samsun-Merkez varieties, is also closely linked in the dendrogram due to their genetic proximity to these two varieties.

3.2. Morphological parameters

When the morphological characteristics of the genotypes were examined (Table 5), it was observed that the fruit diameter, number of fruit lobes, harvest times and especially fruit shapes of Bafra-Merkez and Sivas-Yukarikale Village varieties were similar. These similarities are seen with both genetic similarity and genetic relationship dendrogram. The morphological similarities between Samsun-Carsamba and Aydin-Nazilli, which are genetically close, and Karabuk-Safranbolu and Antalya-Kas are also striking. If it is compared the genetic data as dendrogram and morphological data, it seems that there is no grouping due to morphological parameters.

In this study covering the assessment of genetic diversity on tomato landraces, SSR marker analysis gave a great of information about the distinguishing of landraces. However, Nei

and Shannon indexes were not calculated in this study because of the selection effect by individuals and unbalanced pollination on landraces that will not give safe information on population (Hennink and Zeven, 1990). The genetic diversity on tomato landraces in this study is so important that breeders can find any different genotype material carrying desired genes. Morphological data shows that desired genes can be find in these population making pure lines of genotypes for increasing the similarity within bred lines using molecular markers (Okumus, 2007).

Acknowledgements: This Project was funded by Ondokuz Mayıs University with the number of PYO.ZRT.1904.10.032.

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

Informed consent: This manuscript did not involve human or animal participants; therefore informed consent was not collected.

References

- Benor, S., Zhang, M., Wang, Z., & Zhang, H. (2008). Assessment of genetic variation in tomato (*Solanum lycopersicum* L.) inbred lines using SSR molecular markers. *Journal of Genetics and Genomics*, 35(6), 373-379.
- Botstein, D., White, R. L., Skolnick, M., & Davis, R. W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics*, 32(3), 314.
- Brookfield, J. F. Y. (1996). A simple new method for estimating null allele frequency from heterozygote deficiency. *Molecular Ecology*, 5(3), 453-455.
- Can, H., Kal, U., Ozyigit, I. I., Paksoy, M., & Turkmen, O. (2019). Construction, characteristics and high throughput molecular screening methodologies in some special breeding populations: a horticultural perspective. *Journal of Genetics*, 98(3), 86.
- Castellana, S., Ranzino, L., Beritognolo, I., Cherubini, M., Luneia, R., Villani, F., & Mattioni, C. (2020). Genetic characterization and molecular fingerprint of traditional Umbrian tomato (*Solanum lycopersicum* L.) landraces through SSR markers and application for varietal identification. *Genetic Resources and Crop Evolution*, 67(7), 1807-1820.
- Cox, S. (2000). From discovery to modern commercialism: the complete story behind *Lycopersicon esculentum*, Colorado State, USA. <http://www.landscapeimagery.com/articles.html>, Last accessed on July 18, 2021.
- Dilbirli, E., (2007). Bitkisel biyolojik çeşitlilik ve genetik kaynaklarının sürdürülebilir kullanım stratejilerinin değerlendirilmesi. *Ankara Üniversitesi, Fen Bilimleri Enstitüsü, Doktora Tezi*.
- Doyle, J. J., & Doyle, J. L. (1991). Isolation of plant DNA from fresh tissue plant. *Plant Molecular Biology*, 9, 340-342.
- Fentik, D. A. (2017). Review on genetics and breeding of tomato (*Lycopersicon esculentum* Mill). *Advances in Crop Science and Technology*, 5(5), 1-6.
- He, C., Poysa, V., & Yu, K. (2003). Development and characterization of simple sequence repeat (SSR) markers and their use in determining relationships among *Lycopersicon esculentum* cultivars. *Theoretical and Applied Genetics*, 106(2), 363-373.
- Hennink, S., & Zeven, A. C. (1990). The interpretation of Nei and Shannon-Weaver within population variation indices. *Euphytica*, 51(3), 235-240.
- Henareh, M., Dursun, A., & Mandoulakani, B. A. (2015). Genetic diversity in tomato landraces collected from Turkey and Iran revealed by morphological characters. *Acta Scientiarum Polonorum Hortorum Cultus*, 14(2), 87-96.
- Henareh, M., Dursun, A., Abdollahi-Mandoulakani, B., & Haliloğlu, K. (2016). Assessment of genetic diversity in tomato landraces using ISSR markers. *Genetika*, 48(1), 25-35.
- Hocaoglu-Ozyigit, A., Ucar, B., Altay, V., & Ozyigit, I. I. (2020). Genetic diversity and phylogenetic analyses of Turkish cotton (*Gossypium hirsutum* L.) lines using ISSR markers and chloroplast trnL-F regions. *Journal of Natural Fibers*, 1-14.
- Kaemmer, D., Weising, K., Beyermann, B., Borner, T., Eppel, J. T., & Kahlm, G. (1995). Oligonucleotide fingerprinting of tomato DNA. *Plant Breeding*, 114(1), 12-17.
- Garcia-Martinez, S., Andreani, L., Garcia-Gusano, M., Geuna, F., & Ruiz, J. J. (2006). Evaluation of amplified fragment length polymorphism and simple sequence repeats for tomato germplasm fingerprinting: utility for grouping closely related traditional cultivars. *Genome*, 49(6), 648-656.
- Nei, M. (1972). Genetic distance between populations. *The American Naturalist*, 106(949), 283-292.
- Nei, M., & Li, W. H. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences*, 76(10), 5269-5273.
- Nei, M. (1987). *Molecular evolutionary genetics*. (pp. 1-512). Columbia University Press, New York.
- Okumus, A. (2007). Genetic variation and relationship between Turkish flint maize landraces by RAPD markers. *American Journal of Agricultural and Biological Sciences*, 2(2), 49-53.
- Osma, E., Ozyigit, I. I., Leblebici, Z., Demir, G., & Serin, M. (2012). Determination of heavy metal concentrations in tomato (*Lycopersicon esculentum* Miller) grown in different station types. *Romanian Biotechnological Letters*, 17(1), 6962-6974.
- Ozgen, M., Adak, S., Karagoz, A. Ulukan, H. (1995). Bitkisel gen kaynaklarının korunma ve kullanımı. *Türkiye Ziraat Mühendisliği 4. Teknik Kongresi, Ankara, Ziraat Bankası Kültür Yayınları*, 26, 309-343.
- Peakall, R., & Smouse, P. (2010). GenA1Ex 6.4. Appendix 1: Methods and statistics. <http://wenku.baidu.com/view/f483ee3243323968011c927c.html>, Last accessed on July 18, 2021.
- Rohlf, F. J. (1998). NTSyS-p.c. Numerical Taxonomy and Multivariate Analysis System (Version 2.0). Exeter Software Publishers Ltd., Setauket.
- Schouten, H. J., Tikunov, Y., Verkerke, W., Finkers, R., Bovy, A., Bai, Y., & Visser, R. G. (2019). Breeding has increased the diversity of cultivated tomato in The Netherlands. *Frontiers in Plant Science*, 10, 1606.
- Sefc, K. M., Lefort, F., Grando, M. S., Scott, K. D., Steinkellner, H., & Thomas, M. R. (2001). Microsatellite markers for grapevine: a state of the art. In: Roubelakis-Angelakis, K. A., (eds) *Molecular Biology & Biotechnology of the Grapevine* (pp. 433-463). Springer, Dordrecht.
- Sol Genomics, (2021). The Official Website of Sol Genomics Network, https://solgenomics.net/about/tomato_project/nsf_9872617/index.pl,

- Last accessed on July 18, 2021.
- Tam, S. M., Mhiri, C., Vogelaar, A., Kerkveld, M., Pearce, S. R., & Grandbastien, M. A. (2005). Comparative analyses of genetic diversities within tomato and pepper collections detected by retrotransposon-based SSAP, AFLP and SSR. *Theoretical and Applied Genetics*, 110(5), 819-831.
- Tridge, (2021). The Official Website of Global Sourcing Hub of Food & Agriculture, <https://www.tridge.com/intelligences/tomato/production>, Last accessed on July 18, 2021.
- Vural, H., Esiyok, D., & Duman, I. (2000). *Kültür sebzeleri (sebze yetiştirme)*. (pp. 1-440). Ege Üniv., Ziraat Fak., Bornova, İzmir.
- Wien, H. C. (1987). *The physiology of vegetable crops*. (pp. 1-207). Cornell University, Ithaca N.Y., USA.
- World's Top Exports, (2021). The Official Website of World's Top Exports, <http://www.worldstopexports.com/tomatoes-exports-country>, Last accessed on July 18, 2021.
- Yazgan, A., & Fidan, S. (1996). Tokat koşullarına uygun kiraz domates (*Lycopersicon esculentum* Mill. var. *cerasiforme*) çeşitlerinin belirlenmesi. *GAP I. Sebze Tarimi Sempozyumu, Sanliurfa*, 19-23.

Cite as: Okumus, A., & Dagidir, S. (2021). Assessment of genetic diversity on tomato (*Lycopersicon esculentum* Mill.) landraces using SSR molecular markers in Turkey. *Front Life Sci RT*, 2(2), 51-59.