Research Paper / 115-125

Molecular diversity and genetic structure of rainfed durum wheat genotypes using SCoT markers

Zahra Moradi Kheibari[1], Reza Azizinezhad[1*], Ali Mehras Mehrabi[2], Mahmood Khosrowshahli[1], Alireza Etminan[2]

[1]Department of Biotechnology and Plant Breeding, Science and Research Branch, Islamic Azad University, P. O. Box: 31749-55433, Tehran, Iran.
[2]Department of Biotechnology and Plant Breeding, Kermanshah Branch, Islamic Azad University, Kermanshah, Iran.
[*]Corresponding author, Email: r.azizi@srbiau.ac.ir. Tel: +98-26-32815803.

Received: 08 Nov 2021; Accepted: 12 Apr 2022.
DOI: 10.30479/IJGPB.2022.16475.1307

Abstract

Selection- and conservation-based breeding programs require the study of genetic diversity. In this study, a collection of durum wheat consisting of 90 rainfed genotypes was subjected to the analysis of genetic diversity and population structure based on polymorphisms obtained from the Start Codon Targeted (SCoT) marker system. Out of 26 initial primers tested, 15 primers produced scorable polymorphism and were therefore, selected for further analyses. On average, 11.27 polymorphic fragments were observed for each primer per reaction. Polymorphism Information Content (PIC) ranged from 0.10 to 0.32 per locus with an average of 0.23 per primer. Resolving power (Rp) was varied from 0.98 to 5.80. The structure analysis classified the assessed population into 3 subpopulations. Besides, the Neighbor-Joining phylogenetic tree and Principal Coordinate Analysis separated genotypes into 3 and 5 distinct clusters. The Analysis of Molecular Variance (AMOVA) revealed high intra-population diversity. The gene flow index (Nm) indicated a relatively small probability of gene flow between the studied subsets. The Nei's gene diversity (n), Shannon's information index (I), and allele distribution statistics revealed that the individuals of subpopulation-2 had a significant capacity for genetic diversity. In conclusion, the studied SCoT primers had a high discriminating power and therefore, were efficient for evaluating genetic diversity in the durum wheat. The results of this study revealed the existence of a significant genetic diversity between the studied genotypes. Besides, the individuals of subpopulation-2 had a notable level of genetic diversity that can be used for various breeding purposes.

Key words: Gene flow, Inter-population differentiation, Neighbor-Joining algorithm, Principal Coordinate Analysis.

INTRODUCTION

Durum wheat (*Triticum turgidum* L. subsp. *durum*) is a remarkable species of cereal family that has a high economic value. Durum is the raw material of semolina and some other local products (Romano *et al.,* 2021). The worldwide demand for durum wheat is gradually increasing, and statistics show that universal production of durum wheat in the 2016-2017 season reached 40.7 m t (Beres *et al.,* 2020). In Iran, wheat is cultivated under irrigated and rainfed conditions. In addition, approximately two-thirds of the area under wheat cultivation is devoted to rainfed cultivation. However, it accounts for one-third of wheat production (Mohammadi and Amri, 2013).

Genetic diversity is an essential element of selection-based genetic and breeding programs (Sharma *et al.,* 2021). Preliminary studies of diversity were mainly based on morpho-physiological traits and protein markers (Khang *et al.,* 2021). However, they were not

IJGPB is an open access journal under the CC BY license (http://creativecommons.org/licenses/BY/4.0/)

entirely successful due to environmental influences. With the advent of DNA molecular markers, they have become an essential tool in studying genetic diversity as they are not affected by environmental factors (Nadeem *et al.,* 2018).

There are many reports for genotyping durum wheat and other wild and cultivated wheat using high throughput DNA markers (Lotti *et al.,* 2000; Soleimani *et al.,* 2002; Medini *et al.,* 2005; Karaca, 2008).

The Start codon target polymorphism (SCoT) is a simple but powerful DNA marking method proposed by Collard and Mackill (2008). In this method, primers which are the fundamental component of the PCR reaction are designed based on a short conserved region near the ATG translation initiation codon in plant genes in both DNA strands.

Heidari *et al.* (2017) screened a set of 17 durum wheat genotypes using 14 SCoT markers and reported that the primers generated 99 polymorphic bands with an average of 7.07 bands per primer. The values of Marker Index (MI) and Polymorphism Information Content (PIC) indicated that SCoT markers had a high efficiency in detecting genetic variation in durum wheat. In another study, Etminan *et al.* (2016) investigated genetic variation among 43 durum wheat genotypes using six SCoT markers. Based on the levels of polymorphisms, they concluded that the SCoT marker system was a valuable tool for the detection of variation among durum wheat genotypes.

Genetic erosion due to severe cultivation of modified cultivars, as well as consecutive breeding cycles, has narrowed the genetic background of the cultivated genotypes. As a result, the remaining diversity in the gene pool may be insufficient for future breeding programs. Hence, the continuous evaluation of the genetic diversity of the current plant germplasm is necessary for breeding-based programs. Therefore, the present paper aimed to study the diversity and genetic structure of a population of rainfed durum wheat comprising of 90 genotypes using SCoT marker-based polymorphism.

MATERIALS AND METHODS

Plant materials and DNA extraction

In this study, a collection consisting of 90 rainfed durum wheat genotypes maintained at Kermanshah Dryland Research Center, Iran, was investigated (Table 1). This collection contains an important part of the rainfed durum wheat genotypes stored in the center, which performed well in various experiments. The CTAB method suggested by Murray and Thompson (1980) was used to extract the genomic DNA followed by a quality test over a 1% agarose gel.

Table 1. Name, code, and the origin of the 90 genotypes of rainfed durum wheat used in the study.

Genotype	Name/Code	Origin	Genotype	Name/Code	Origin
G1	CDSS06B00053S-099Y-099M-12Y-2B-04Y-0B	CIMMYT	G12	CDSS09Y00310S-099Y-034M-12Y-0M-04Y-0B	CIMMYT
G2	CDSS09Y00029S-099Y-020M-9Y-0M-04Y-0B	CIMMYT	G13	CDSS09Y00318S-099Y-014M-27Y-0M-04Y-0B	CIMMYT
G3	CDSS09Y00241S-099Y-022M-10Y-0M-04Y-0B	CIMMYT	G14	CDSS09Y00843T-099Y-034M-9Y-0M-04Y-0B	CIMMYT
G4	CDSS09Y00286S-099Y-026M-24Y-0M-04Y-0B	CIMMYT	G15	CDSS09Y00211S-099Y-041M-11Y-0M-04Y-0B	CIMMYT
G5	CDSS09Y00762T-099Y-024M-20Y-0M-04Y-0B	CIMMYT	G16	CDSS09Y00211S-099Y-041M-16Y-0M-04Y-0B	CIMMYT
G6	CDSS10Y00498T-099Y-018M-12Y-1M-06Y-0B	CIMMYT	G17	CDSS09Y00314S-099Y-029M-24Y-0M-04Y-0B	CIMMYT
G7	CDSS10Y00498T-099Y-018M-18Y-1M-06Y-0B	CIMMYT	G18	CDSS09Y00327S-099Y-041M-19Y-0M-04Y-0B	CIMMYT
G8	CDSS10Y00504T-099Y-037M-10Y-1M-06Y-0B	CIMMYT	G19	CDSS09Y00762T-099Y-024M-19Y-0M-04Y-0B	CIMMYT
G9	CDSS09B00165S-099Y-010M-4Y-3M-06Y-0B	CIMMYT	G20	CDSS09Y00805T-099Y-09M-5Y-0M-04Y-0B	CIMMYT
G10	CDSS09B00171S-099Y-041M-1Y-3M-06Y-0B	CIMMYT	G21	CDSS08B00131T-099Y-027M-6Y-0M-04Y-0B	CIMMYT
G11	CGSS02Y00004S-2F1-6Y-0B-1Y-0B	CIMMYT	G22	CDSS05B00007S-6Y-0M-1Y-4M-0Y	CIMMYT

Table 1 (continued). Name, code, and the origin of the 90 genotypes of rainfed durum wheat used in the study.

Genotype	Name/Code	Origin	Genotype	Name/Code	Origin
G23	CDSS06Y00326S-44Y-0M-5Y-1M-0Y	CIMMYT	G51	CMSS08B01011S-099B-099Y-41B-0Y	CIMMYT
G24	CDSS07Y00768D-3B-01Y-03M-6Y-1B-04Y-0B	CIMMYT	G52	CMSS09Y01198T-099TOPB-099Y-099B-108Y-0Y	CIMMYT
G25	CDSS08Y00760T-0TOPB-099Y-08M-12Y-1M-0Y	CIMMYT	G53	CMSS09Y01199T-099TOPB-099Y-099B-45Y-0Y	CIMMYT
G26	CDSS04B00362T-0TOPY-16Y-0M-1Y-0M-2Y-0B	CIMMYT	G54	CMSS09Y01200T-099TOPB-099Y-099B-57Y-0Y	CIMMYT
G27	CDSS06Y00646T-0TOPB-24Y-0M-4Y-1M-0Y	CIMMYT	G55	CMSS09Y01201T-099TOPB-099Y-099B-18Y-0Y	CIMMYT
G28	CDSS06Y00625T-0TOPB-34Y-0M-2Y-1M-0Y	CIMMYT	G56	CMSS09Y01201T-099TOPB-099Y-099B-27Y-0Y	CIMMYT
G29	CDSS06Y00497S-28Y-0M-4Y-4M-0Y	CIMMYT	G57	CMSS09Y01201T-099TOPB-099Y-099B-49Y-0Y	CIMMYT
G30	CDSS06Y00816T-0TOPB-61Y-0M-8Y-1M-0Y	CIMMYT	G58	CMSS09Y01202T-099TOPB-099Y-099B-83Y-0Y	CIMMYT
G31	CDSS06B00472T-099Y-099M-11Y-4M-04Y-0B	CIMMYT	G59	CMSS09Y01202T-099TOPB-099Y-099B-102Y-0Y	CIMMYT
G32	CDSS06B00488T-099Y-099M-5Y-3M-04Y-0B	CIMMYT	G60	CMSS09Y01203T-099TOPB-099Y-099B-5Y-0Y	CIMMYT
G33	CDSS07Y00544T-099Y-099M-15Y-0M-04Y-0B	CIMMYT	G61	CMSS09Y01203T-099TOPB-099Y-099B-10Y-0Y	CIMMYT
G34	CDSS07Y00544T-099Y-099M-24Y-3M-04Y-0B	CIMMYT	G62	CMSS09Y01203T-099TOPB-099Y-099B-35Y-0Y	CIMMYT
G35	CDSS06Y00816T-0TOPB-61Y-0M-1Y-4M-0Y	CIMMYT	G63	CMSS09Y01203T-099TOPB-099Y-099B-39Y-0Y	CIMMYT
G36	CDSS06Y00674T-0TOPB-4Y-0M-3Y-4M-0Y	CIMMYT	G64	CMSS09Y01203T-099TOPB-099Y-099B-49Y-0Y	CIMMYT
G37	CDSS07B00338S-099Y-013M-4Y-1M-0Y	CIMMYT	G65	CMSS09Y01203T-099TOPB-099Y-099B-54Y-0Y	CIMMYT
G38	CDIB02Y00011T-B-4B-3Y-3B-3Y-2B-1Y-2B-2Y-1B-0Y	CIMMYT	G66	CMSS09Y01203T-099TOPB-099Y-099B-58Y-0Y	CIMMYT
G39	CDSS09Y00415S-099Y-021M-2Y-0M-04Y-0B	CIMMYT	G67	CMSS09Y01203T-099TOPB-099Y-099B-72Y-0Y	CIMMYT
G40	CMSS08B01003S-099B-099Y-45B-0Y	CIMMYT	G68	CMSS09Y01203T-099TOPB-099Y-099B-82Y-0Y	CIMMYT
G41	CMSS08B00996S-099B-099Y-36B-0Y	CIMMYT	G69	CMSS09Y01203T-099TOPB-099Y-099B-86Y-0Y	CIMMYT
G42	CMSS08B01001S-099B-099Y-38B-0Y	CIMMYT	G70	CMSS09Y01203T-099TOPB-099Y-099B-104Y-0Y	CIMMYT
G43	CMSS08B01001S-099B-099Y-40B-0Y	CIMMYT	G71	CMSS09Y01204T-099TOPB-099Y-099B-21Y-0Y	CIMMYT
G44	CMSS08B01003S-099B-099Y-45B-0Y	CIMMYT	G72	CMSS09Y01204T-099TOPB-099Y-099B-42Y-0Y	CIMMYT
G45	CMSS08B01004S-099B-099Y-29B-0Y	CIMMYT	G73	CMSS09Y01204T-099TOPB-099Y-099B-50Y-0Y	CIMMYT
G46	CMSS08B01009S-099B-099Y-2B-0Y	CIMMYT	G74	CMSS09Y01204T-099TOPB-099Y-099B-61Y-0Y	CIMMYT
G47	CMSS08B01009S-099B-099Y-7B-0Y	CIMMYT	G75	CMSS09Y01204T-099TOPB-099Y-099B-63Y-0Y	CIMMYT
G48	CMSS08B01011S-099B-099Y-20B-0Y	CIMMYT	G76	CMSS09Y01204T-099TOPB-099Y-099B-65Y-0Y	CIMMYT
G49	CMSS08B01011S-099B-099Y-25B-0Y	CIMMYT	G77	CMSS09Y01204T-099TOPB-099Y-099B-91Y-0Y	CIMMYT
G50	CMSS08B01011S-099B-099Y-30B-0Y	CIMMYT	G78	CMSS09Y01204T-099TOPB-099Y-099B-92Y-0Y	CIMMYT

Table 1 (continued). Name, code, and the origin of the 90 genotypes of rainfed durum wheat used in the study.

Genotype	Name/Code	Origin	Genotype	Name/Code	Origin
G79	CMSS09Y01205T-099TOPB-099Y-099B-26Y-0Y	CIMMYT	G85	CMSS09Y01209T-099TOPB-099Y-099B-155Y-0Y	CIMMYT
G80	CMSS09Y01205T-099TOPB-099Y-099B-61Y-0Y	CIMMYT	G86	IRD2010-11-003-OMAR-OMAR-OSAR-OSAR-OSAR-1SAR	Iran
G81	CMSS09Y01205T-099TOPB-099Y-099B-65Y-0Y	CIMMYT	G87	Saji	Iran
G82	CMSS09Y01205T-099TOPB-099Y-099B-72Y-0Y	CIMMYT	G88	Zahab	Iran
G83	CMSS09Y01205T-099TOPB-099Y-099B-77Y-0Y	CIMMYT	G89	SRN-1/KILL//2*FOLTA-1	CIMMYT
G84	CMSS09Y01205T-099TOPB-099Y-099B-89Y-0Y	CIMMYT	G90	Imren	Turkey

CIMMYT*:* The International Maize and Wheat Improvement Center.

Genotyping

In this study, we used 15 SCoT primers, previously developed by Collard and Mackill (2008) (Table 2). The PCR was performed in a volume of 20 µl, consisting of 2 µl template DNA (25 ng μL^{-1}), 12.6 µl double distilled water, 2 µl 1X buffer, 1.5 µl $MgCl_2$ (50 mM), 0.4 µl dNTP (10 mM), 1.2 µl primer, 0.3 µl Taq DNA Polymerase. The PCR reaction program was set as follows: an initial denaturation for 5 min at 94 °C, followed by 45 cycles comprising of denaturation (94 °C for 45 s); a variable melting temperature (from 52 - 60 °C for 45 s for each primer presented in Table 2); primer elongation for 90 s at 72 °C; and a final extension for 10 min at 72 °C. The PCR products were then electrophoresed on a 1.5% agarose gel and the gel images were prepared after staining under UV light.

Data analysis

First, the information obtained on the gels was converted into zero (absence) and one (presence) to form a data matrix. Then, some genetic parameters were estimated based on the obtained data matrix. These were total amplified fragments (TAB), the number of polymorphic bands (NPB), and the percentage of polymorphic bands (PPB). The Polymorphic Information Content (PIC) was calculated as $PIC=1-\sum p_i^2$, where p_i is the *i*th allele frequency (Serrote *et al.,* 2020). The Marker Index (MI) criterion was calculated as $MI=PIC\times EMR$, where EMR (Effective Multiplex Ratio) was defined as $EMR=n\times\beta$ where $\beta=NPB\times(NPB/TAB)$ (Kumar and Agrawal, 2019). The resolving power (Rp) of markers was estimated as $R_p=\sum I_b$, where $I_b=\sum 1-(2\times|0.5-p_i|)$ where p_i is the proportion of genotypes containing the band (1) (Kumar and Agrawal, 2019). The genetic variability estimates such as the effective number of alleles (Ne), Shannon's Information Index (I), Nei's gene diversity (h), the gene flow criterion (Nm), and diversity among subpopulations (Gst) were determined with POPGENE software (Negisho *et al.,* 2021). The principal coordinate analysis (PCoA) and Molecular Analysis of Variance (AMOVA) were performed using the GenAlex 6.5 software (Peakall and Smouse, 2012). The subpopulations of genetically similar individuals were recognized using the STRUCTURE software (version 2.3.4). The resulting data were then transferred into the Structure Harvester software (Earl and vonHoldt, 2012) to detect the optimal number of subpopulations according to Evanno *et al.* (2005). In addition, genotypes were clustered based on the neighbor-joining method and Jaccard's distance matrix using MEGA 6.0 software.

RESULTS

Polymorphism and discriminating power of the SCoT markers

PCR amplification profiles, produced by four typical SCoT primers (SCoT02, SCoT21, SCoT24, and SCoT26), were shown in Figure 1. Also, Table 2 presents all data regarding polymorphism obtained, as well as discriminating power statistics calculated for each SCoT primer. A total of 8317 amplicons were obtained, with 554.5 amplicons per primer. Generally, a total of 169 polymorphic bands (PB) were obtained. On average, each primer had 11.27 PB. The lowest and highest number of PB were found for primers SCoT21 and SCoT26, respectively. Average Polymorphic Information Content (PIC) values over loci were varied from 0.10 (SCoT07) to 0.32 (SCoT21) with an average of 0.23 per primer.

Table 2. Name, sequences, melting temperature, discriminating power statistics, and the polymorphisms obtained from 15 SCoT primers on the 90 genotypes of rainfed durum wheat.

Primer	Sequence (5'→3')	Tm	TAB	NPB	PPB%	PIC	MI	EMR	Rp
SCoT01	CAACAATGGCTACCACCA	49.1	10	10	100	0.13	0.87	6.67	1.62
SCoT02	CAACAATGGCTACCACCC	53.7	10	10	100	0.22	1.47	6.67	3.13
SCoT03	CAACAATGGCTACCACCG	53.7	15	15	100	0.27	4.05	15.00	5.71
SCoT04	CAACAATGGCTACCACCT	49.1	12	12	100	0.24	2.30	9.60	3.58
SCoT05	CAACAATGGCTACCACGA	49.1	12	11	91.67	0.30	2.18	8.07	4.60
SCoT07	CAACAATGGCTACCACGG	53.7	12	9	75	0.10	0.38	5.40	0.98
SCoT08	CAACAATGGCTACCACGT	49.1	11	9	81.82	0.27	1.19	5.40	3.53
SCoT09	CAACAATGGCTACCAGCA	49.1	13	13	100	0.29	3.27	11.27	5.22
SCoT11	AAGCAATGGCTACCACCA	49.1	12	9	75	0.20	0.81	5.40	2.47
SCoT12	ACGACATGGCGACCAACG	51.4	10	10	100	0.21	1.40	6.67	2.60
SCoT14	ACGACATGGCGACCACCG	56	13	11	84.62	0.24	1.69	8.07	3.60
SCoT15	ACGACATGGCGACCGCGA	56	15	15	100	0.14	2.10	15.00	2.73
SCoT21	CACCATGGCTACCACCAT	53.7	9	8	88.89	0.32	1.20	4.27	3.73
SCoT24	CCATGGCTACCACCGCCA	565	11	11	100	0.21	1.69	8.07	3.20
SCoT26	ACAATGGCTACCACCATC	49.1	16	16	100	0.26	4.44	17.07	5.80
Mean			12.07	11.27	93.13	0.23	1.94	8.84	3.50

Tm: Melting temperature, TAB: Total amplified bands, NPB: The number of polymorphic bands, PPB: Percentage of polymorphic bands, Ne: The number of effective alleles, h: Nei's gene diversity, I: Shannon's information index, PIC: The average polymorphism information content values for each primer, MI: Marker index, Rp: Resolving power.

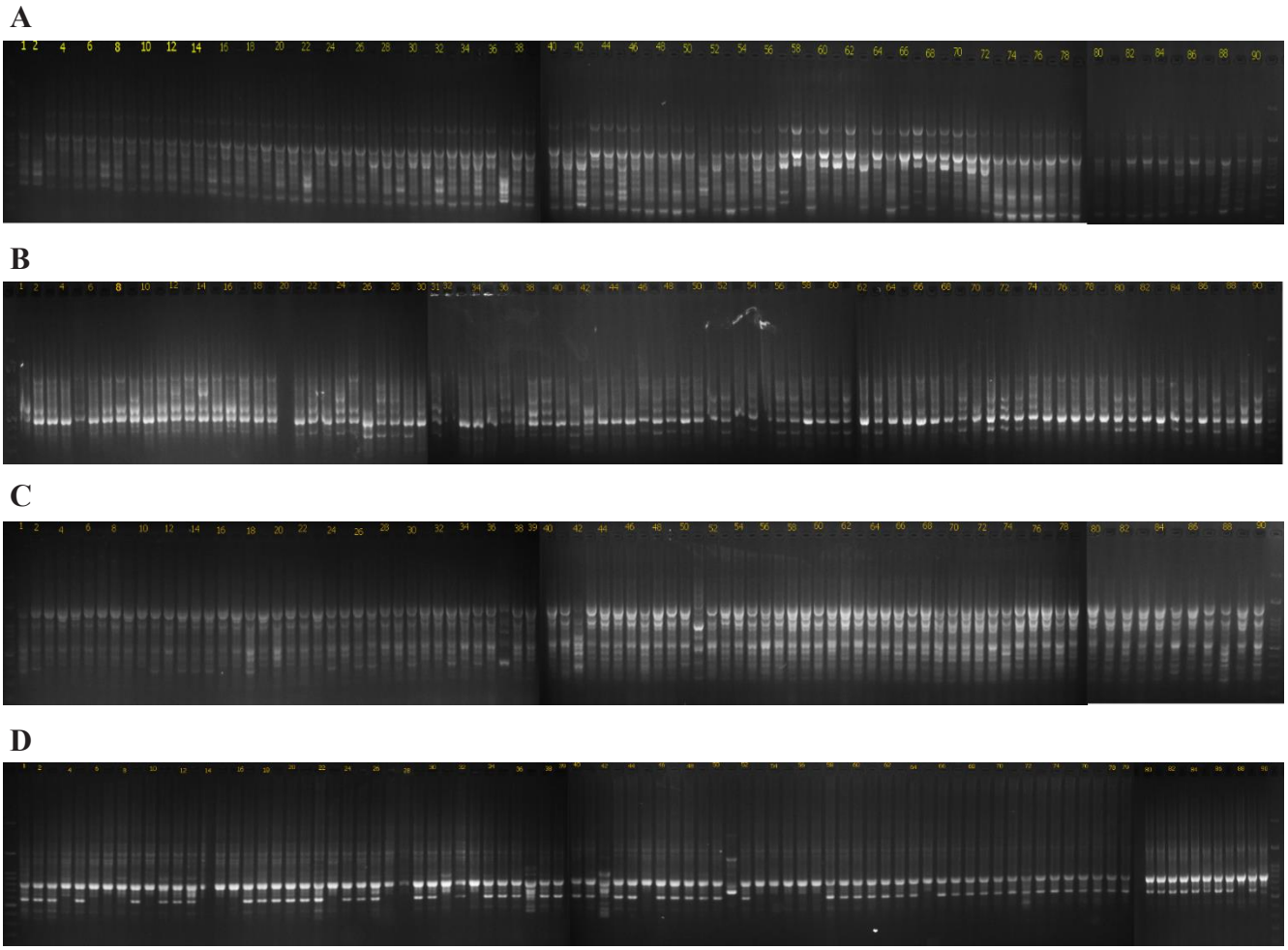

Figure 1. Amplification profile of some SCoT primers in 90 rainfed durum wheat genotypes used in the study: **A**, **B**, **C**, and **D:** show the PCR amplification induced by SCoT02, SCoT21, SCoT24, and SCoT26 primers.

Marker Index (MI) fluctuated from 0.38 (SCoT07) to 4.44 (SCoT26) with an average of 1.94. The average effective multiplex ratio (EMR) value was calculated to be 8.84, ranging from 4.27 for SCoT21 to 17.07 for SCoT26. The highest and lowest Resolving power (Rp) index were observed for SCoT26 (5.80) and SCoT07 (0.98), respectively, with an average of 3.50.

Classification of primers based on their Rp and PIC values presented in Table 2 revealed that SCoT03, SCoT04, SCoT05, SCoT08, SCoT09, SCoT14, SCoT21, and SCoT26 markers performed well in both statistics unlike the other group, which had low values for both measurements. i.e., SCoT01, SCoT07, SCoT15.

Structure and cluster analyses

First, the value of ΔK (likelihood values of partitioning) was plotted against the different values of K (the number of subpopulations) to find the maximum value of ΔK (Figure 2). From Figure 2, the value of ΔK was at its maximum with 3 subpopulations (K=3). Accordingly, the genotypes were divided into three groups containing 31, 53, and 6 members, respectively (Figure 3). Most of the genotypes introduced from CIMMYT were categorized in groups 1 and 2, while Zahab and Imren genotypes were clustered together in the third group (Figure 3).

The lowest genetic distance (Jaccard coefficient) was observed between G71 and G73 (0.09) whereas G1 and G51 with 0.72 had the highest distance (Figure 4). Therefore, it can be concluded that the primers used had good potential for calculating genetic distances as well as discovering associations between durum

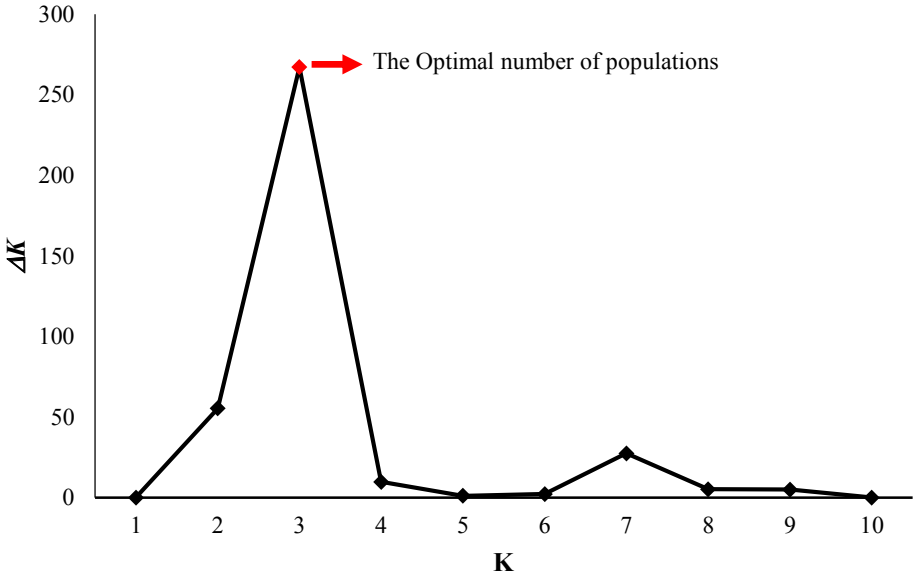

Figure 2. The optimal number of subpopulations of the 90 genotypes of rainfed durum wheat based on molecular data obtained from 15 SCoT primers.

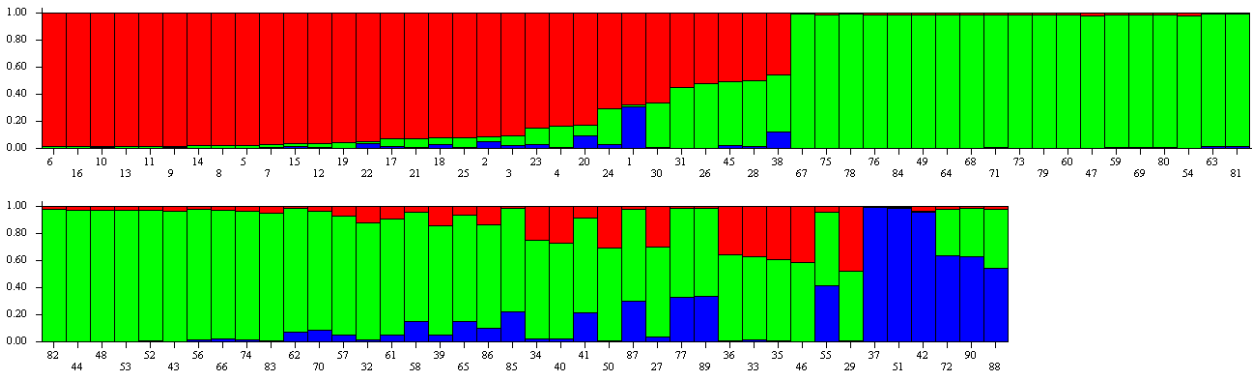

Figure 3. Population structure of the 90 genotypes of rainfed durum wheat based on Bayesian model obtained from 15 SCoT primers polymorphism.

genotypes. Neighbor-joining cluster analysis separated the 90 durum wheat genotypes into three clusters (Figure 4). The third cluster was further classified into clusters 3-1 and 3-2. Cluster 1 consisted of G51, G37, and G42 genotypes with different banding patterns compared to other genotypes.

The first to third principal coordinates (PCoA) accounted for 26.68, 11.88, and 5.91 percent of the total variation (Figure 5). As shown in Figure 5, the 90 durum genotypes were separated into five groups.

Genetic diversity among subpopulations

Criteria for the diversity of the three identified subpopulations have been listed in Table 3. The subpopulations had a percentage of polymorphic loci (PPL) of 65.75, 89.50, and 41.44%, respectively. Subpopulation-2 had the highest number of different alleles (Na=1.87) compared to the lowest Na (1.13) as showed by subpopulation-3. The highest number of effective alleles belonged to subpopulation-2 (1.31), whereas subpopulation-3 with 1.27 had the lowest value. Also, Shannon's Information Index values were found to be 0.28, 0.32, and 0.23 for the subpopulations, respectively. The Fixation index (Fst) values were found to be 0.4951, 0.5323, and 0.0389 for subpopulations 1 to 3, respectively. Similarly, Table 3 shows that the h values were 0.18, 0.20, and 0.19, respectively. From the results, it can be deduced that genotypes grouped in subpopulation-2 had the utmost genetic dissimilarity than other subpopulations, while genotypes in subpopulation-3 had a higher genetic similarity.

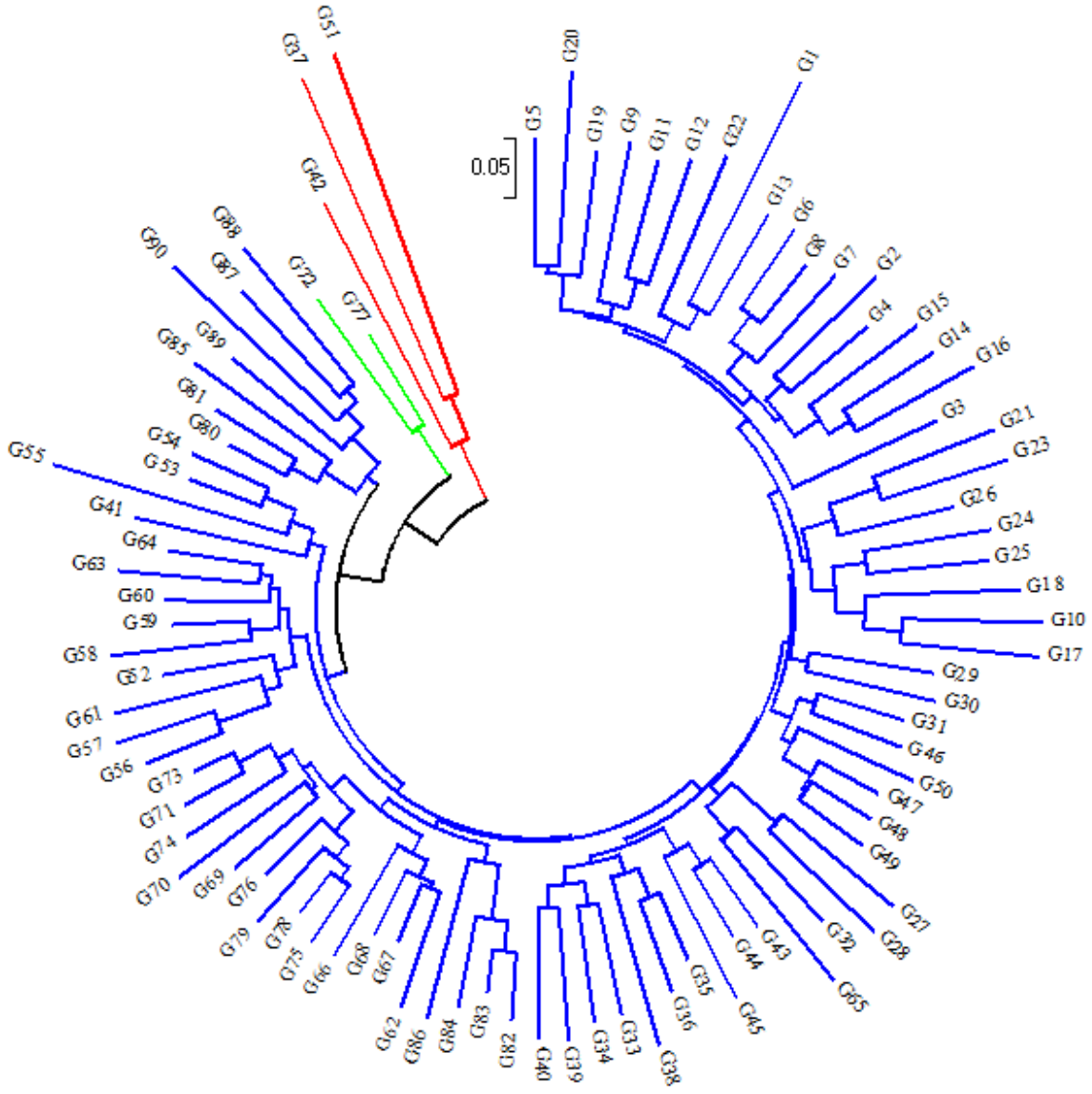

Figure 4. Cluster analysis of the 90 genotypes of rainfed durum wheat according to the Neighbor-joining method and Jaccard's distance coefficient obtained from 15 SCoT primers polymorphism.

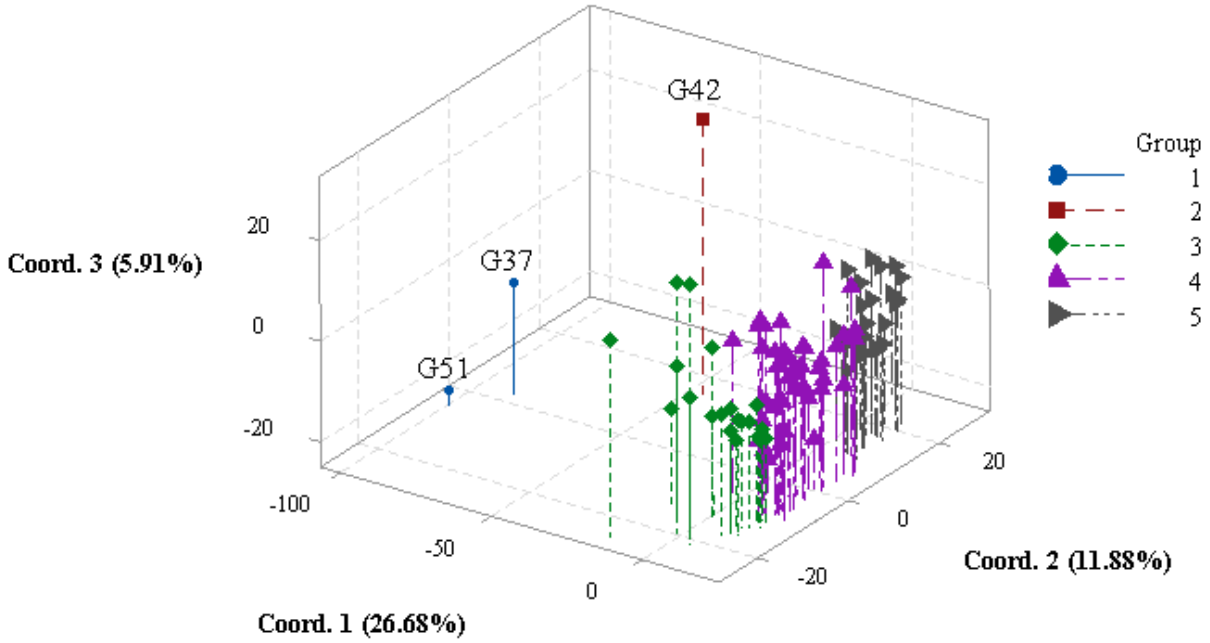

Figure 5. Principal coordinate analysis based on polymorphisms obtained from 15 SCoT primers on the 90 genotypes of rainfed durum wheat.

Table 3. Summary of genetic variation statistics for the three subpopulations of rainfed durum wheat genotypes estimated using 15 SCoT markers.

Subpopulation	Size	PPL%	Na	Ne	I	h	F_{ST}
1	31	65.75	1.5±0.06	1.28±0.02	0.28±0.02	0.18±0.01	0.4951
2	53	89.50	1.87±0.03	1.31±0.02	0.32±0.02	0.2±0.01	0.5323
3	6	41.44	1.13±0.06	1.27±0.03	0.23±0.02	0.16±0.01	0.0389

PPL%: The percentage of polymorphic loci, Na: The number of different alleles, Ne: The number of effective alleles, I: Shannon's Information index, h: Nei's gene diversity, F_{ST}: Fixation index.

The distribution statistics of alleles in the three subpopulations resulting from the STRUCTURE analysis was estimated (Data not shown). According to the results, the number of different bands ranged from 130 for subpopulation-3 to 176 for Subpopulation-2. There were 5 and 21 bands unique to subpopulation-1 and subpopulation-2, respectively, whereas no unique band was found for subpopulation-3. Also, the maximum number of different bands with a frequency of≥5% was found for subpopulation-2. Generally, no locally common bands were found in 50% or fewer populations. Besides, subpopulation-3 had the lowest mean diversity (0.157) whereas subpopulation-2 had the highest one (0.2).

Based on the AMOVA, variation within the populations accounted for 85% of the total variation (P<0.001). In contrast, the variation between the populations accounted for only 15% of the total variation. Therefore, most of the amplicons had useful information for identifying subgroups (Table 4). Besides, Table 4 shows that the Gst and Nm values were 0.22 and 1.75, respectively.

DISCUSSION

Investigation of genetic diversity among rainfed wheat genotypes is an effective strategy to identify unique populations with specific applications for use in short-term and long-term purposes of plant breeding. Results of this study suggested that the SCoT marker system, with an average polymorphism of 93.13%, was able to detect a significant level of genetic diversity among rainfed durum wheat genotypes. The percentage of polymorphisms observed in this study was significantly higher compared to that reported by Karaca (2008) for ISSR and RAPD primers.

Table 4. Analysis of molecular variance of the three subpopulations of rainfed durum wheat based on molecular data obtained from 15 SCoT primers.

c	df	SS	MS	Est. Var.	Variation (%)	Gst	Nm	Phi	P(rand>=data)
Among Pops	2	187.788	93.894	3.192	15				
Within Pops	87	1543.645	17.743	17.743	85				
Total	89	1731.433		20.935	100	0.22	1.75	0.152	0.001

df: Degree of freedom, SS: Sum of squares, MS: Means of squares, Est. Var: Estimated variance components, Gst: Inter-population differentiation, Nm: Estimate of gene flow.

Parameters such as PIC have been usually used to evaluate the informative potential of SCoT markers (Heidari *et al.,* 2017). In this work, the PIC values were low to moderate. Despite some consistencies, the PIC values were lower compared to those reported by Heidari *et al.* (2017) and Etminan *et al.* (2016). This difference could not be due to the number of alleles amplified by SCoT primers because the primers used in this study, which were similar to the primers used by Heidari *et al.* (2017) and Etminan *et al.* (2016) amplified the same number of alleles. Therefore, the discrepancy can be due to the difference in the frequency of alleles for each locus which in turn was due to differences in genotypes and population size. Also, in this research, the population size was much larger (90 genotypes) compared to the populations studied by Heidari *et al.* (2017) and Etminan *et al.* (2016) (17 and 43 genotypes, respectively). As shown by Iles *et al.* (2003) the larger the population, the more accurate the estimate of allele frequency. Accordingly, our calculation of allele frequencies for each locus may be closer to reality resulting in more reliable PIC values.

MI, EMR, and Rp values have been used to evaluate the discriminatory power of different molecular marker systems (Amom *et al.,* 2020). In this study, the average values obtained for MI, EMR, and Rp were equal to 1.94, 8.84, and 3.5, respectively. These results were consistent with those reported by Etminan *et al.* (2016). Also, the Rp values were higher compared to the report of Heidari *et al.* (2017). The scatter chart of markers based on their PIC and Rp values identified a group of markers including SCoT03, SCoT04, SCoT05, SCoT08, SCoT09, SCoT14, SCoT21, and SCoT26 that performed well in both statistics. Therefore, these primers should be considered in future molecular studies on durum wheat. In contrast, SCoT01, SCoT07, and SCoT15 had low values for both criteria.

In this study, the classification of genotypes using PCoA (Figure 5) was consistent with that in the neighbor-joining (NJ) phylogenetic tree (Figure 4), and the structure analysis pattern (Figure 3) with G42, G51, and G37 separated from the remaining populations. Nevertheless, there were some inconsistencies. As shown in the three above figures, some individuals with the same structure (Figure 3) have been classified into different clusters by NJ (Figure 4) and PCoA. (Figure 5). Such differences can be attributed to the mathematical model used in the above three grouping methods. The mathematical model used by STRUCTURE software was designed to sort individuals into Hardy–Weinberg populations while the neighbor-joining clustering is a distance-based technique and uses the star decomposition method.

In this study, the AMOVA revealed that the estimated variance (Est. Var) within populations was 5.6 times the estimated variance among populations (Table 4). Also, the inter-population differentiation (Gst) value was assessed to be 0.22, indicating that 22% of the total genetic variability was among the three subpopulations while 88% of the total genetic variability was within subpopulations. Therefore, the results obtained from AMOVA and the estimation of Gst were consistent. Also, the Gst criterion was greater than 0.15 showing that there was a relatively high genetic distinction between subpopulations (Nei, 1973). Fixation index (F_{ST}) shows the degree of gene differentiation among populations in terms of allele frequencies. When F_{ST} is less than 0.05 then the genetic differentiation is small while, an F_{ST} between 0.05 to 0.15 shows a moderate genetic differentiation. In addition, when F_{ST} is 0.15 to 0.25 then the genetic differentiation is large and an F_{ST} greater than 0.25 shows a very large genetic differentiation (Meirmans and Hedrick, 2011). Hence, sub-populations 1 and 2, (with F_{ST} values of 0.4951, and 0.5323, respectively) had a very large genetic differentiation whereas, the genetic differentiation of sub-population 3 was small. (F_{ST}=0. 0389).

Nei (h) and Shannon (I) genetic diversity indices are among the most appropriate genetic parameters that have been used in various studies to determine the diversity and distinction among and within

subpopulations. High values of these parameters indicate high genetic diversity. h measures the average genetic diversity per locus. In other word, it is the probability that, at a single locus, any two alleles, chosen at random from the population, are different to each other. Therefore, the average h over all loci is an estimate of the extent of genetic variability in the population (Hennink and Zeven, 1990). Accordingly, similar to that revealed by the F_{ST} coefficient, subpopulations 2, 1 and 3 with h index of 0.2, 0.18, and 0.16 showed the highest genetic diversity, respectively (Table 3). Moreover, in the present study, the gene flow (Nm) among subpopulations was recorded to be 1.75, which indicates a relatively small probability of gene flow between the studied subsets. As demonstrated by Varvio *et al.* (1986) The Nm index is defined as the number of migrants per subpopulation in one generation and values higher than one indicate minor genetic differentiation among subpopulations. Furthermore, allele distribution statistics showed that subpopulation-2 had the highest number of different bands (176), the highest number of unique bands (21), the maximum number of different bands with a frequency of≥5%, and the highest mean diversity (0.2) compared to the two other subpopulations. These results indicated that the individuals of subpopulation-2 had a significant level of genetic diversity that can be used for various breeding purposes. Also, the high values of the percentage of polymorphic loci (89.5%), a notable number of different alleles (1.87), and the highest number of effective alleles (1.31) parameters confirmed this conclusion. In agreement with our findings, Jlassi *et al.* (2021) using SSR; Shaygan *et al.* (2021) using CBDP and ISSR; Alemu *et al.* (2020) using SNP marker techniques reported a notable amount of genetic diversity in durum genotypes.

CONCLUSIONS

According to the results, the informativeness parameters showed an appropriate level of discriminating power of the SCoT primers. Therefore, the amplification fragments of this marker have high efficiency for the analysis of genetic diversity among durum wheat germplasm. Results also showed a remarkable level of genetic diversity among studied durum wheat genotypes. The AMOVA revealed that intra-population diversity was responsible for the highest genetic variations, while inter-population differentiation accounted for only 15% of total genetic variations suggesting that genotypes had a wide genetic differentiation. The Nei's gene diversity, Shannon's information index, and allele distribution statistics revealed that the individuals of subpopulation-2 had a considerable level of diversity that can be used for various breeding purposes.

ACKNOWLEDGMENTS

The authors thank the Sararud Dryland Agricultural Research Institute of Kermanshah for its sincere cooperation in providing the genotypes.

REFERENCES

Alemu A., Feyissa T., Letta T., and Abeyo B. (2020). Genetic diversity and population structure analysis based on the high density SNP markers in Ethiopian durum wheat (*Triticum turgidum* ssp. *durum*). *BMC Genetics,* 21: 18. DOI: https://doi.org/10.1186/s12863-020-0825-x.

Amom T., Tikendra L., Apana N., Goutam M., Sonia P., Koijam A. S., Potshangbam A. M., Rahaman H., and Nongdam P. (2020). Efficiency of RAPD, ISSR, iPBS, SCoT and phytochemical markers in the genetic relationship study of five native and economical important bamboos of North-East India. *Phytochemistry,* 174: 112330.

Beres B. L., Rahmani E., Clarke J. M., Grassini P., Pozniak C. J., Geddes C. M., Porker K. D., May W. E., and Ransom J. K. (2020). A systematic review of durum wheat: Enhancing production systems by exploring genotype, environment, and management (G×E×M) Synergies. *Frontiers in Plant Science,* 11: 568657.

Collard B. C. Y., and Mackill D. J. (2008). Start codon Targeted (SCoT) polymorphism: a simple, novel DNA marker technique for generating gene-targeted markers in plants. *Plant Molecular Biology Reporter,* 27: 86. DOI: https://doi.org/10.1007/s11105-008-0060-5.

Earl D. A., and vonHoldt B. M. (2012). Structure harvester: a website and program for visualizing structure output and implementing the Evanno method. *Conservation Genetics Resources,* 4: 359–361.

Etminan A., Pour-Aboughadareh A., Mohammadi R., Ahmadi-Rad A., Noori A., Mahdavian Z., and Moradi Z. (2016). Applicability of start codon targeted (SCoT) and inter-simple sequence repeat (ISSR) markers for genetic diversity analysis in durum wheat genotypes. *Biotechnology & Biotechnological Equipment,* 30: 1075–1081.

Evanno G., Regnaut S., and Goudet J. (2005). Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology,* 14: 2611–20.

Heidari P., Etminan A., Azizinezhad R., and Khosroshahli M. (2017). Genomic variation studies in durum wheat (*Triticum turgidum* ssp. *durum*) using CBDP, SCoT and ISSR markers. *Indian Journal* of *Genetics* and Plant Breeding*,* 77: 379–386.

Hennink S., and Zeven A. (1990). The interpretation of Nei and Shannon-Weaver within population variation

indices. *Euphytica,* 51: 235–240.

Iles M. M., Walters K., and Cannings C. (2003). Recombination can evolve in large finite populations given selection on sufficient loci. *Genetics,* 165: 2249–2258.

Jlassi I., Bnejdi F., Saadoun M., Hajji A., Mansouri D., Ben-Attia M., El-Gazzah M., and El-Bok S. (2021). SSR markers and seed quality traits revealed genetic diversity in durum wheat (*Triticum durum* Desf.). *Molecular Biology Reports,* 48: 3185–3193.

Karaca M. (2008). Comparative analysis of genetic diversity in Turkish durum wheat cultivars using RAPD and ISSR markers. *Journal of Food, Agriculture & Environment,* 6(3-4): 219–225.

Khang N. H. M., Quang N. T., Mai H. N. X., Phuong N. D. N., Thao N. P., and Quoc N. B. (2022). Genetic characterization of coconut (*Cocos nucifera* L.) varieties conserved in Vietnam through SCoT marker-based polymorphisms. *Genetic Resources and Crop Evolution*, 69: 385–398. DOI: https://doi.org/10.1007/s10722-021-01237-x.

Kumar J., and Agrawal V. (2019). Assessment of genetic diversity, population structure and sex identification in dioecious crop, *Trichosanthes dioica* employing ISSR, SCoT and SRAP markers. *Heliyon,* 5: e01346-e01346.

Lotti C., Salvi S., Pasqualone A., Tuberosa R., and Blanco A. (2000). Integration of AFLP markers into an RFLP-based map of durum wheat. *Plant Breeding,* 119: 393–401.

Medini M., Hamza, S., Rebai, A., and Baum, M. (2005). Analysis of genetic diversity in Tunisian durum wheat cultivars and related wild species by SSR and AFLP markers. *Genetic Resources and Crop Evolution,* 52: 21–31.

Meirmans P. G., and Hedrick P. W. (2011). Assessing population structure: FST and related measures. *Molecular Ecology Resources,* 11: 5–18.

Mohammadi R., and Amri A. (2013). Genotype×environment interaction and genetic improvement for yield and yield stability of rainfed durum wheat in Iran. *Euphytica,* 192: 227–249.

Murray M. G., and Thompson W. F. (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Research,* 8: 4321–5.

Nadeem M. A., Nawaz M. A., Shahid M. Q., Doğan Y., Comertpay G., Yıldız M., Hatipoğlu R., Ahmad F., Alsaleh A., Labhane N., Özkan H., Chung G., and Baloch F. S. (2018). DNA molecular markers in plant breeding: current status and recent advancements in genomic selection and genome editing. *Biotechnology & Biotechnological Equipment,* 32: 261–285.

Negisho K., Shibru S., Pillen K., Ordon F., and Wehner G. (2021). Genetic diversity of Ethiopian durum wheat landraces. *Plos one,* 16: e0247016.

Nei M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences,* 70(12): 3321–3. DOI: 10.1073/pnas.70.12.3321.

Nouraein M., Khavari-Khorasani S., and Akhavan M. (2020). Screening cumin (*Cuminum cyminum* L.) landraces for resistance to Fusarium oxysporum f. sp. cumini. *Australasian Plant Pathology,* 49: 295–305.

Peakall R., and Smouse P. E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research--an update. *Bioinformatics (Oxford, England),* 28: 2537–2539.

Romano A., Ferranti P., Gallo V., and Masi P. (2021). New ingredients and alternatives to durum wheat semolina for a high quality dried pasta. *Current Opinion in Food Science*, 41: 249–259.

Serrote C. M. L., Reiniger L. R. S., Silva K. B., Rabaiolli S. M. d. S., and Stefanel C. M. (2020). Determining the polymorphism information content of a molecular marker. *Gene,* 726: 144175.

Sharma P., Mehta G., Shefali, Muthusamy S. K., Singh S. K., and Singh G. P. (2021). Development and validation of heat-responsive candidate gene and miRNA gene based SSR markers to analysis genetic diversity in wheat for heat tolerance breeding. *Molecular Biology Reports,* 48: 381–393.

Shaygan N., Etminan A., Majidi Hervan I., Azizinezhad R., and Mohammadi R. (2021). The study of genetic diversity in a minicore collection of durum wheat genotypes using agro-morphological traits and molecular markers. *Cereal Research Communications,* 49: 141–147.

Soleimani V. D., Baum B. R., and Johnson D. A. (2002). AFLP and pedigree-based genetic diversity estimates in modern cultivars of durum wheat [*Triticum turgidum* L. subsp. *durum* (Desf.) Husn.]. *Theoretical and Applied Genetics,* 104: 350–357.

Varvio S.-L., Chakraborty R., and Nei M. (1986). Genetic variation in subdivided populations and conservation genetics. *Heredity,* 57: 189–198.