

## Assessment of Genetic Diversity among Iranian *Triticum* Germplasm Using Agro-morphological Traits and Start Codon Targeted (SCoT) Markers

A. POUR-ABOUGHADAREH<sup>1\*</sup>, J. AHMADI<sup>1</sup>, A.A. MEHRABI<sup>2</sup>, A. ETMINAN<sup>3</sup> and M. MOGHADDAM<sup>4</sup>

<sup>1</sup>Department of Crop Production and Breeding, Imam Khomeini International University, Qazvin, Iran

<sup>2</sup>Department of Agronomy and Plant breeding, University of Ilam, Ilam, Iran

<sup>3</sup>Department of Plant breeding, Kermanshah Branch, Islamic Azad University, Kermanshah, Iran

<sup>4</sup>Department of Plant Breeding and Biotechnology, University of Tabriz, Tabriz, Iran

(Received 10 November 2016; Accepted 31 January 2017;

Communicated by A. Börner)

The knowledge about genetic diversity in the wild relatives of wheat provides useful information for breeding programs and gene pool management. In the present study, an assessment of agro-morphological diversity and molecular variability among 70 accessions of *Triticum*, belonging to *T. boeoticum*, *T. urartu*, *T. durum* and *T. aestivum* species, collected from different regions of Iran was made. According to phenotypic analysis, all traits except peduncle length, stem diameter and the number of seeds per spike indicated a high level of diversity among studied accessions. Also, principal component analysis identified six components that explained 87.53% of the total variation in agro-morphological traits. In molecular analysis, 15 start codon targeted (SCoT) polymorphism primers produced 166 bands, out of which, 162 (97.59%) were polymorphic. Analysis of molecular variance (AMOVA) indicated the 63% of the variation resided among populations. The maximum value of polymorphism information content (*PIC*), the observed (*Na*) and effective (*Ne*) number of alleles, Nie's gene diversity (*He*) and Shannon's information index (*I*) was detected for *T. boeoticum* than the other species. The SCoT-based tree revealed three different groups corresponding to the genomic constitution in *Triticum* germplasm, which was in part confirmed by STRUCTURE and principal coordinate (PCoA) analyses. Our results indicated a remarkable level of genetic diversity among studied Iranian *Triticum* species, especially *T. boeoticum*, which can be of interest for future breeding and other analyses associated with future studies of the wild relatives of wheat. More importantly, our results revealed that SCoT markers could be used to accurately evaluate genetic diversity and phylogenetic relationships among different *Triticum* species.

**Keywords:** *Triticum*, genetic diversity, SCoT, structure analysis

### Introduction

The genus *Triticum* belongs to the tribe Triticeae within the Pooideae subfamily of the grass family Poaceae. This genus is important in the wheat gene pool due to their evolutionarily relationship with the major agricultural crop *T. aestivum* L. van Slageren (1994)

\*Corresponding author; E-mails: a.poraboghdashareh@edu.ikiu.ac.ir; a.poraboghdashareh@gmail.com

described three *Triticum* species including diploid (*T. boeoticum* Bioss., *T. urartu* Thum. ex Gandil. and *T. monococcum* L.), tetraploid (*T. dicoccoides*, *T. araraticum* and *T. turgidum* var. *durum* Desf.) and hexaploid (*T. aestivum* L. and *T. zhukovskii* Men. et Erizan.) species. Eco-geographical regions of Iran, especially in the northwest, west and southwest of country (east of the Fertile Crescent) are main center of distribution of wild relative of wheat, so that compositions of *Aegilops* and *Triticum* species as the richest wheat gene pool has been found in these areas (Salimi et al. 2005). Information regarding the extent and pattern of genetic variation in the crop plants enhance germplasm management and utilization. There are several methods to assess diversity such as molecular, biochemical, morphological and botanical characterization. Of these, the botanical and agromorphological characters are the first and easiest approaches for the assessment of genetic diversity in different plant genera. Many studies showed that phenotypic indicators are still helpful in identification and evaluation of genetic diversity (Salimi et al. 2005; Zhang et al. 2015; Alsaleh et al. 2016). On the other hand, the molecular tools provide good information in studies on genetic diversity and phylogenetic relationships among different plant crops, because they indicate genetic variation in greater detail without interference from environmental conditions. With the advances in plant molecular biology, a number of molecular marker techniques have been developed and used widely in evaluating genetic diversity, population structure, genetic relationship and germplasm management. In recent years, advances in molecular markers have resulted in the use of random DNA markers to functional markers, gene-targeted and the development of novel DNA-based marker systems (Poczai et al. 2013). Collard and Mackill (2009) developed a simple and novel DNA-based marker technique start codon targeted (SCoT) polymorphism. This marker system uses the short conserved region in plant genes surrounding the ATG translation start codon. This technique can provides more information associated with biological characters compared with other DNA marker techniques. As a new molecular marker system, SCoT markers with high polymorphism and high efficiency has been successfully applied in many plant species (Xiong et al. 2011; Gorji et al. 2011; Bhattacharyya et al. 2013; Hamidi et al. 2014; Rajesh et al. 2015; Etminan et al. 2016). In the present study, we aimed to evaluate genetic variability among populations of *Triticum* germplasm using agro-morphological traits and molecular data obtained from SCoT markers. The assessment of the efficiency of SCoT marker technique for diagnostic fingerprinting in wheat germplasm was the other aim of this study.

## Materials and Methods

### *Plant materials*

A total of 70 *Triticum* accessions were collected from natural habitats, deserts, valleys and mountains from the Zagros and central Elburz mountains located from northern, north-western, northeastern to southern, and southwestern regions of Iran. The evaluated accessions consisted of two diploid species (*T. boeoticum* Bioss. and *T. urartu* Thum. ex Gandil.), one tetraploid species (*T. durum* Desf.) and one hexaploid species (*T. aestivum* L.).

Detailed information about genomic constitution and eco-geographical distribution of these materials is listed in Table S1\*.

### *Agro-morphological data*

The field experiment was carried out in Research Station of the Faculty of Agriculture, Department of Agronomy and Plant breeding, University of Ilam, Iran (Latitude 33°39'N, Longitude 46°22'E, and Altitude 1445 m above sea level) during 2014–2015 cropping seasons, and 20 agro-morphological traits were recorded following IBPGR handbook (1985). Augmented design with two blocks was used with four repeated check accessions randomly arranged in each block. Plot size was one row 3 m long with 30 cm spacing between rows. The land was left as fallow in the previous year, and 50 kg h<sup>-1</sup> triple super-phosphate (TSP 46%) fertilizer was added to the soil before planting. The climate is characterized by mean annual precipitation of 244.38 mm per year, and mean annual temperature of 16.8 °C and annual maximum temperature of 22.6 °C. The agro-morphological traits were: (1) number of days-to-booting, (2) to-heading, (3) to-anthesis, (4) to-physiological maturity, (5) grain filling period, (6) plant height, (7) peduncle length, (8) leaf length, (9) leaf width, (10) number of stem nodes, (11) stem diameter, (12) number of fertile tillers per plant, (13) main spike length, (14) number of spikelet, (15) number of seeds per spikelet, (16) number of seeds per spike, (17) 1000-seed weight, (18) seed yield per plant, (19) biomass per plot and (20) harvest index. The traits 6–17 were measured on the basis of 10 randomly chosen plants per plot. To avoid the loss of seed and estimate seed yield per plant in accessions of *T. boeoticum* and *T. urartu*, all spikes were protected with trap. Finally, total above ground dry weight was measured as biomass. Harvest index was determined from the ration of seed yield to biomass. Augmented design analysis and adjusted means was computed following the SAS PROC GLM outlined by Wolfinger et al. (1997). Statistical analysis included: descriptive statistics (i.e.: mean, range, variance, standard division and coefficient of variability) and principal components analysis (PCA). Also, Shannon-Weaver (*H'*/SW) diversity index (Hutchenson 1970) for agro-morphological traits was estimated. This diversity index was classified as high ( $H' \geq 0.60$ ), intermediate ( $0.40 \leq H' \leq 0.60$ ) and low ( $0.10 \leq H' \leq 0.40$ ) as described by Eticha et al. (2005).

### *DNA extraction, PCR amplification and molecular analysis*

From each accession, after the seed germination and growth, the total genomic DNA was isolated from the young leaves of glasshouse-grown plants according to the CTAB protocol (Doyle and Doyle 1987). DNA quality was also analyzed by 0.8% agarose. After optimizing PCR, of the 25 primers tested, the 15 primers that amplified scoreable polymorphic DNA bands were selected for further analysis. The PCRs were carried out in the reaction mixture 20 µL volume, with 10 µL master mix 2XPCR (Ready-to-use PCR master mix 2X), 6 µL ddH<sub>2</sub>O, 2 µL of isolated DNA from each sample and 2 µL of each primer.

\*Further details about the Electronic Supplementary Material (ESM) can be found at the end of the article.

Amplification was run at 94 °C for 5 min, followed by 45 cycles of denaturation at 94 °C for 45 seconds, primer annealing at 56–62.8 °C (varied for each primer) for 45 seconds and primer elongation at 72 °C for 90 seconds. The final extension was 10 min at 72 °C. The amplification reaction products were detected by 1.5% denaturing agarose gels stained with safe view II. Then, we scored all the visible and clear polymorphic bands in the SCoT profile as 0 and 1 on the basis of the absence and presence of the band, respectively. Discriminatory power of the primers was estimated by means of three informativeness parameters; Polymorphism Information Content (PIC), Resolving power (Rp) and Marker Index (MI). The distribution of genetic variation within and among populations (AMOVA) was estimated using GenAlix 6.5 software (Peakall and Smouse 2006). The genetic diversity parameters at species levels, viz. the percentage of polymorphic loci (*PPL*), the observed (*N<sub>a</sub>*) and effective numbers of alleles (*N<sub>e</sub>*), Nei's gene diversity (*N<sub>e</sub>*), Shannon's information index (*I*), inter-population differentiation (*G<sub>st</sub>*) and gene flow (*N<sub>m</sub>*) were estimated using POP-GENE software version 1.31 (Yeh et al. 1997). The Jaccard's similarities matrix from the calculative data was calculated using the NTSYS-pc software (Rohlf 2000). Neighbor joining (NJ) dendrogram was generated using DARwin (Perrier et al. 2003). A principal co-ordinate analysis (PCoA) was performed to analyze the partitioning of genetic variance among the species using GenAlix 6.5. Bayesian clustering model implanted in STRUCTURE software version 2.3.4 (Pritchard et al. 2000) was used for estimating the number of optimal clusters (*K*) and the proportion of assignment of individuals from each of the assumed population to each of the inferred genetic clusters. An admixture model was used to infer the number of '*K*' with prior population information. All analyses were done with a burning time period of 5000 and a Markov Chain Monte Carlo (MCMC) replication number set up to 100,000. The program was run 10 times for each '*K*', ranging from 1 to 10. The optimal *K* value, as true cluster number (Evanno et al. 2005), was determined using STRUCTURE HARVESTER program (Earl and von Holdt 2012).

## Results

### *Phenotypic analysis*

Descriptive statistics for 20 agro-morphological traits are shown in Table 1. Results indicated a large variation in these germplasm based on agro-morphological traits. Among measured characters, the large variability corresponded to biomass followed by plant height, 1000-seed weight, peduncle length, days-to-heading, harvest index and number of seeds per spike. The highest CV value was observed for seed yield, peduncle length, leaf length, biomass, leaf width, harvest index, and 1000-seed weight, respectively. In contrast, developmental characters, the number of stem nodes and the number of spikelets per spike showed the lowest variation. Also, Shannon-Weaver diversity index (*H'*SW) ranged from 0.49 (for days-to-heading) to 0.99 (for main spike length) with an average 0.71. Most traits showed relatively high levels of diversity, so that the highest value of this index recorded for main spike length, the number of spikelets per spike, leaf width, the

Table 1. Mean, maximum, minimum, variance, standard deviation (SD) and coefficient of the variation (CV) values for each characteristic evaluated for *Triticum* accessions

Character	Mean	Min	Max	SD	CV	H <sup>2</sup> SW <sup>a</sup>
Days-to-booting	157.93	139.63	170.38	8.08	5.11	0.75
Days-to-heading	165.80	149.37	188.62	9.18	5.54	0.82
Days-to-anthesis	170.63	160.25	195.75	8.22	4.81	0.78
Days-to-physiological maturity	195.18	176.13	212.13	7.77	3.98	0.89
Grain filling period	24.55	9.63	35.38	5.31	21.62	0.80
Plant height (cm)	71.95	43.02	118.48	14.16	19.67	0.83
Peduncle length (cm)	22.08	5.61	76.63	12.47	56.50	0.34
Number of stem nodes	3.41	2.98	5.03	0.54	15.81	0.66
Stem diameter (cm)	2.94	1.99	4.82	0.66	22.64	0.59
Leaf length (cm)	11.36	3.50	24.06	6.19	54.44	0.96
Leaf width (cm)	0.85	0.31	1.99	0.41	47.65	0.61
Number of fertile tillers per plant	8.62	2.25	18.25	3.18	36.89	0.80
Spike length (cm)	9.28	4.88	13.80	1.76	18.92	0.67
Number of spikelets per spike	18.48	11.73	27.07	3.33	18.00	0.89
Number of seeds per spikelet	1.91	0.85	2.95	0.52	27.43	0.59
Number of seeds per spike	34.62	14.30	59.90	8.87	25.62	0.85
1000-seed weight (g)	28.03	12.12	57.38	12.74	45.44	0.83
Seed yield per plant (g)	8.01	0.33	22.08	4.64	57.89	0.95
Biomass (g)	38.73	10.82	91.18	19.23	49.66	0.91
Harvest index	19.84	2.80	46.22	9.10	45.84	0.92

<sup>a</sup> Shannon-Weaver diversity index (Hutchenson 1970).

number of grains per spike, leaf length, stem diameter, the number of fertile tillers per plant, plant height, days-to-booting and to-physiological maturity, respectively. Taking into account of agro-morphological means comparison we identified the best species for each character (Table S2). For instance, the shorter time of development such as days-to-heading, to-anthesis and to-maturity followed by leaf length and leaf width belonged to the *T. durum* species. *T. urartu* and *T. boeoticum* indicated the lowest value for grain filling period. Also, the highest value of plant height, peduncle length and the number of seeds per spike recorded for *T. urartu* followed by *T. boeoticum* and *T. aestivum*. In the case of biomass and harvest index characters, the highest amount was showed for *T. durum* and *T. aestivum*. Principal component analysis (PCA) was performed due to the reduction of data to break down the total variance in the main features and a few numbers of the independent new components. The factor loading, eigenvalue communality and estimated specific variance for the studied traits are shown in Table S3. For the 70 accessions, the top six PCAs explained 87.53% of the total variation. The first component justified

39.33% of the total variation, and was strongly influenced by days-to-booting, to-heading, to-anthesis, the number of stem nodes, leaf length, leaf width, stem diameter, the number of spikelets, the number of seeds per spikelet and seeds per spike, 1000-seed weight, seed yield, biomass and harvest index. The second component accounted for 19.76% of the total variation and was mainly explained by days-to-booting, to-heading, to-anthesis, to-maturity, as well as plant and peduncle length. The third component justified 8.80% of the total variation and was strongly influenced by grain filling period, the number of fertile tillers per plant and harvest index. The fourth, fifth and sixth components accounted for 8.10, 6.47 and 5.05% of the total variation and were influenced by the number of spikelet, the number of seeds per spike and plant height, respectively.

### *Molecular analysis*

Out of 25 primers tested, 15 primers were selected, which generated in total 166 bands, out of which, 162 (97.59%) were polymorphic. The number of total bands ranged from 8 (SCoT-2 and SCoT-5) to 13 (SCoT-12, SCoT-15, SCoT-19 and SCoT-20), with an average of 11.06 bands per primer combination. The number of polymorphic bands varied from 8 (SCoT-2, SCoT-5 and SCoT-21) to 13 (SCoT-12 and SCoT-19). The percentage of polymorphic bands across the primers ranged from 80% (SCoT-21) to 100% (SCoT-2, SCoT-3, SCoT-5, SCoT-6, SCoT-7, SCoT-12, SCoT-14, SCoT-18 and SCoT-19). The mean of polymorphic information content (*PIC*) values for the primers was 0.476 and primers SCoT1 showed the highest value (0.50), while primer SCoT 2 had the lowest (0.371). The primers SCoT-19 and SCoT-21 had the highest and lowest marker index (*MI*) value, respectively. Resolving power (*R<sub>p</sub>*) for primers ranged from 4.93 to 12, with an average 9.29, which primers SCoT-19 and SCoT-3 revealed the highest and lowest value, respectively (Table 2). The AMOVA results (Table S4) showed that 63% genetic variation within populations, while the variation between the populations was 10% ( $\Phi_{PT} = 0.369$ ;  $P = 0.010$ ). This result indicates that 63 percent of amplified fragments are informative for separating the species from each other, and also suggesting that the majority of variation found is within species groups, rather than between species groups. In addition, the results of population analysis showed that the diversity among populations (*G<sub>st</sub>*) and gene flow (*N<sub>m</sub>*) was 0.23 and 3.04, respectively. A summary of the genetic diversity parameters of the four populations (Table S5) indicates that the observed number of alleles (*N<sub>a</sub>*) ranged from 1.41 (*T. urartu*) to 1.72 (*T. boeoticum*) with a mean of 1.58. The means of *He*, *Ne* and *I* indices were most for *T. boeoticum* (*He* = 0.27; *Ne* = 1.46; *I* = 0.41) than the other species. In contrast, the lowest values of these parameters were estimated for *T. urartu* (*He* = 0.24; *Ne* = 1.41; *I* = 0.36). The highest percentage of polymorphic loci (*PPL*) was found among *T. boeoticum* accessions (83.73 %) and the lowest one (65.66 %) among the accessions of *T. urartu*.

In order to explore relationships among the studied accessions, Jaccard's similarity coefficients was calculated according to binary data matrix of amplified bands for the 70 *Triticum* accessions (coefficient matrix not shown). Pairwise genetic similarity ranged from 0.19 to 0.81. The accessions IUGB-00516 (*T. aestivum*) and IUGB-00911 (*T. aesti-*

Table 2. Start Codon Targeted (SCoT) primers and their amplification results generated in 70 accessions of *Triticum* using 15 primers

Primer	Sequence	Ta	TAB	NPB	PPB	PIC	Rp	MI
SCoT-2	CAACAATGGCTACCAACC	56	8	8	100	0.499	7.32	3.99
SCoT-3	CAACAATGGCTACCAACCG	56	10	10	100	0.371	4.93	3.72
SCoT-5	CAACAATGGCTACCAACGA	53.7	8	8	100	0.500	8.11	3.99
SCoT-6	CAACAATGGCTACCAACGC	56	11	11	100	0.432	6.93	4.75
SCoT-7	CAACAATGGCTACCAACGC	56	12	12	100	0.492	10.51	5.91
SCoT-12	ACGACATGGCGACCAACG	58.2	13	13	100	0.465	9.55	6.04
SCoT-13	ACGACATGGCGACCATCG	58.2	11	10	90	0.465	8.08	4.23
SCoT-14	ACGACATGGCGACCAACGC	60.5	11	11	100	0.498	11.61	5.48
SCoT-15	ACGACGTGGCGACCGCGA	60.5	13	12	92.30	0.461	9.38	5.11
SCoT-16	CCATGGCTACCAACCGGCC	62.8	12	11	91.66	0.498	11.18	5.02
SCoT-17	CATGGCTACCAACCGGCC	53	11	10	90.90	0.499	11.46	4.54
SCoT-18	ACCATGGCTACCAACCGCC	60.5	11	11	100	0.499	10.53	5.49
SCoT-19	GCAACAATGGCTACCAACC	56	13	13	100	0.497	12.00	6.46
SCoT-20	AACCATGGCTACCAACGC	58.2	13	12	92.30	0.491	10.37	6.38
SCoT-21	CACCATGGCTACCAACCAT	56	10	8	80	0.469	7.49	2.99
		Mean	11.06	10.73	95.81	0.47	9.29	4.94

Ta temperature annealing, TAB total amplified bands, NPB number of polymorphic, PPB percentage of polymorphism, PIC polymorphism information content, Rp resolving power, MI marker index.



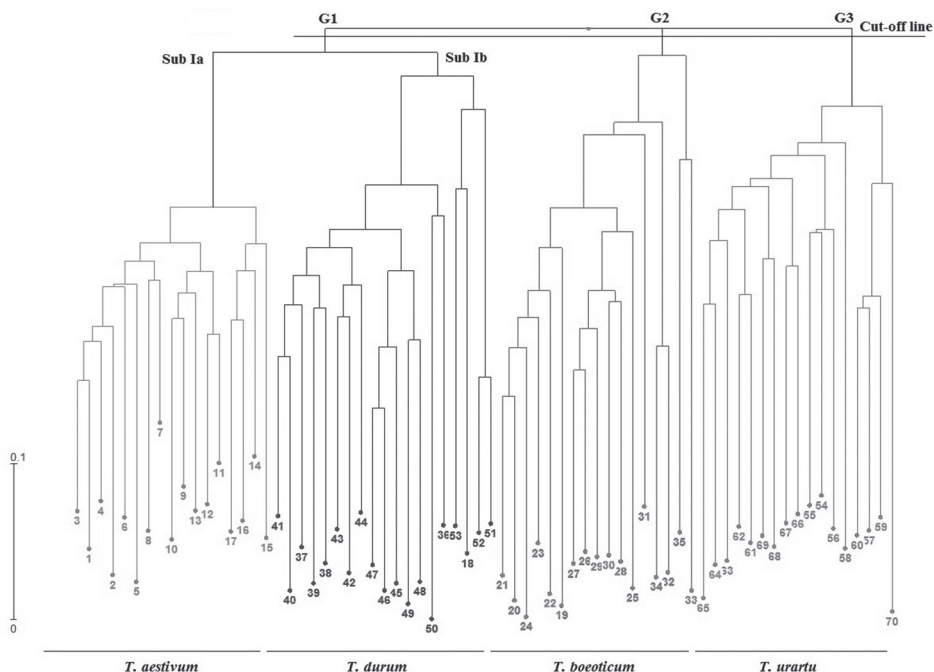


Figure 1. Neighbor-joining tree of start codon targeted (SCoT) polymorphism markers data in the studied *Triticum* species (for explanation of accessions codes, see Table S1)

vum) had the highest genetic similarity (0.81), while the lowest similarity was observed for IUGB-00520 (*T. durum*) with IUGB-00003 and IUGB-00009 (*T. boeoticum*) (0.19). These results recommend that SCoT markers are acceptable marker system for the determination of genetic diversity of *Triticum* germplasm. The neighbor-joining dendrogram grouped the 70 accessions into three main groups (Fig. 1). The first cluster (C1) divided into two sub-clusters (Ia and Ib). Sub-clusters Ia and Ib were consisted of all *T. aestivum* and *T. durum* accessions, respectively. All accessions of *T. boeoticum* and *T. urartu* were clustered into the second (C2) and third (C3) clusters, respectively. Furthermore, principal coordinate analysis (PCoA) showed that, the first two components accounted for 57.54% of the total variation. Two-dimensional plot of PCoA divided the 70 accessions into four distinct groups and its results were in part confirmed by the results of cluster analysis (Fig. S1). The number of genetically distinct populations ( $K$ ; as chosen by Evanno's  $\Delta K$  method) was estimated using the STRUCTURE software to investigative population structure and degree of admixture. The results of Bayesian clustering demonstrated that the true value of  $K$  among the 70 accessions was  $K = 4$  (Fig. S2A). All the assumed four populations were assigned to four distinct genetic groups: (1) *T. boeoticum*, (2) *T. urartu*, (3) *T. durum*, and (4) *T. aestivum* (Fig. S2B). The average membership participation of accessions of afore-said species to their respective inferred clusters ranged from 85 to 97% (Fig. S2C).



## Discussion

In the present study, high levels of agro-morphological diversity were found. Peduncle length, leaf length, leaf width, 1000-seed weight, seed yield, biomass and harvest index characters were highly variable, for which the CV was, > 45% (Table 1). Also, the Shannon-Weaver ( $H'SW$ ) index was selected as the best parameter to estimate phenotypic diversity. The value of the  $H'SW$  revealed the considerable level of diversity, so that according to the classification described by Eticha et al. (2005) all traits except peduncle length, stem diameter and the number of seeds per spike indicated a high level of diversity (> 0.80). The variation exhibited by this germplasm in developmental and quantitative characters indicates that selection for several of these characters might be effective (Moghaddam et al. 1997). In the case of developmental characters there was no significant difference between different species. However, the result showed that the accessions from *T. urartu* and *T. boeoticum* tend to be late in heading, anthesis and booting as compared to the other species. In contrast, *T. urartu* species was found to have shorter grain filling period ranging from 32 to 34 days (Table S2). In Mediterranean dry area as well as similar origins, early heading and early maturity types are suitable because they provide the especial conditions for plants to escape the terminal drought stress that is usually occurred in these regions (Rawashdeh et al. 2007). The length of peduncle is coincident with plant height. Börner et al. (1996) noted plant height and peduncle length are important in disease escape and breeding for resistance to head diseases, because plants with shorter high and length of peduncles are more susceptible. In the case of these characters, accessions of *T. boeoticum* had the highest values of plant height and peduncle length. In this study, high variation in the number of fertile tillers per plant and number of grains per spike were found which, provide the opportunity for breeders to select accessions with high number of spikes and grains per spike (Table S2). This also has been found by Pathak and Nema (1985) in India, who reported that selection for these trait might be effective to improve the landraces of Indian wheat. However, it should be mentioned that, increasing tillering capacity may be an undesirable character in dry climates, because it might result in fast soil moisture depletion, which may be required later in the season, at the critical stages of crop development (Hurd 1971). Also, different levels of variation were observed for spike and its related characters and biomass in the present study. Among the different species, *T. boeoticum* and *T. urartu*, as closer relatives of common wheat, represent a considerable potential for spike length, the number of spiklets and seeds per spike and biomass yield per plant traits. Hence, these species can be attractive for breeding programs (Table S2). Our results show a high level of genetic variability and will help the future breeding programs to achieve ideotypes that are potentially suitable for improvement wheat. Therefore, phenotypic measurements still have the advantage of providing a direct tool in the field for evaluating genetic diversity, select and conserve germplasm. The PCA analysis also showed that, the first six components accounted for 87.53% of total variation among accessions (Table S3) and most traits were consistently present in the first and second components. Therefore, traits related to these components more contribute in the total genetic variation. Similarly, Zhang et al. (2015) demonstrated

that botanical and agro-morphological characters are useful in identification and assessment of genetic diversity in wheat germplasm.

In the current work, we also found a remarkable variation among *Triticum* populations using SCoT marker system. The SCoT primers detected 10.73 polymorphic bands per primer with an average polymorphism of 97.59%, which was relatively high compared to earlier studies using other molecular marker systems such as RAPD, AFLP (Naghavi et al. 2010), microsatellite markers (Moradkhani et al. 2012, 2015) and ISSR (Pour-Aboughadareh et al. 2017). The marker informativeness indices, viz. *R<sub>p</sub>*, *MI* and *PIC* revealed good consent of discriminating power of these primers, suggesting a high efficiency of this DNA-marker to discover genetic diversity among wheat germplasm (Table 2). Of these, *PIC* provides a scale, which helps to determine the effectiveness and potential of the primers used in the fingerprinting process (Powell et al. 1996). Our results showed the *PIC* values ranged from 0.37 to 0.50, with an average of 0.47, which pose the good capability of SCoT primers to the assessment of genotyping in wheat germplasm. The SCoT markers are sensitive to low levels of genetic variations and consequently provide a very useful tool for analyzing population genetics on a wide range of plants as well as recognizing population or species of the same species (Collard and Mackill 2009). In agreement with our results, Gorji et al. (2011) and Rajesh et al. (2015) demonstrated the usefulness of SCoT markers for estimation of genetic diversity fingerprinting of genotypes was more than other DNA markers techniques. Results of analysis of molecular variance (AMOVA) revealed a higher distribution of genetic variation within *Triticum* species (63%) as compared to between species (37%). This is also supported by inter-population differentiation (*G<sub>st</sub>*) value as a suitable parameter for the genetic variation. According to Nei (1978), our populations have been estimated to have the high value of genetic differentiation coefficient (*G<sub>st</sub>*) (0.23). In this respect, populations with *G<sub>st</sub>* values > 0.15, between 0.05 and 0.15 and < 0.05 have the high, medium and low levels of genetic differentiation, respectively (Hamrick 1991). The number of migrants per generation (*N<sub>m</sub>*) is an estimated value from *G<sub>st</sub>* to measure the gene flow, the higher its value, the less genetic differentiation among populations. According to the study conducted by Wright (1951), if *N<sub>m</sub>* > 1, it can inhibit the differentiation among populations due to genetic drift. On the other hand, if *N<sub>m</sub>* < 1, local populations tend to differentiate. In the present study, the amount of *N<sub>m</sub>* among studied populations was 3.04 (Table S4), indicating the possibility of minor gene flow between the studied wild relatives of wheat. In fact, the distribution of genetic variation between and within populations is a function of the rate of gene flow between them, and the extent of gene flow depends on the size and the degree of isolation of its populations as well as the movement seeds or pollens between populations (Dumolin-Lapegue et al. 1997). Besides, based on the genetic variation parameters (Table S5), the highest values of *N<sub>e</sub>*, *N<sub>a</sub>*, *H*, *PPL* indices were estimated for *T. boeoticum* accessions. These results indicate that *T. boeoticum* species as a closer wild relative to the putative donor of A-genome (*T. urartu*) and common wheat may be a good source of genetic diversity to be discovered in crosses with elite wheat.

According to clustering pattern obtained by cluster analysis (Fig. 1) and PCoA plot (Fig. S1), the genetic diversity did not agree with the geographical distribution, as it can

be seen that clustering of individuals was confirmed with the genomic constitution. As shown in Fig. 1, the grouping of *T. aestivum* and *T. durum* species together reveals that these species due to A and B genomes are similar to each other, while *T. boeoticum* and *T. urartu* that carrying A genome were grouped into two distinct clusters. This is also supported by the genetic clusters inferred from STRUCTURE analysis, wherein different accessions grouped into four distinct clusters (Fig. S2). The situation of *T. boeoticum* and *T. urartu* accessions near each other in all dendrograms could explain the local gene flow among populations of different species caused by occasional outcrossing. *T. boeoticum* and *T. urartu* as diploid wheat with A genome are usually difficult to be recognized by visual inspection because these two species are botanical very similar to each other. However, there are several minor morphological and botanical characters that allow us to identify and separate them from each other. For example *T. boeoticum* has long hairs on the veins and little short hairs on the leaf whereas close short hairs cover the leaf blade of *T. urartu*. Other characters in distinguishing the two species from each other are; size difference in anther and awn, number and size differences in seeds in each spikelet, and seed colour (Salimi et al. 2005).

In conclusion, agro-morphological genetic diversity observed in the 70 *Triticum* accessions, which originated from different eco-geographical areas of Iran, indicated that these materials are rich in unexploited potentially useful agro-morphological characters. Moreover, in regard to our finding that showing remarkable levels of genetic diversity between and within Iranian *Triticum* species, it can be suggested that discovering this highly diverse gene pool may result in searching suitable alleles for researchers in wheat evolution, adaptation and improvement. Furthermore, in the present study, we demonstrated the efficiency of SCoT markers for analyzing the genetic diversity as well as genetic relationships among different *Triticum* species. Hence, this marker technique could either be used individually or in combination with other markers techniques to evaluate genetic variation of wheat germplasm and to obtain reliable information about population structure, which would help strategies for effective collection of wheat germplasm and their conservation in genebanks.

## References

- Alsaleh, A., Shehzad Baloch, F., Nachit, M., Ozkan, H. 2016. Phenotypic and genotypic intra-diversity among Anatolian durum wheat “Kundur” landraces. *Biochem. Syst. Ecol.* **65**:9–16.
- Bhattacharyya, P., Kumaria, S., Kumar, S., Tandon, P. 2013. Start codon targeted (SCoT) marker reveals genetic diversity of *Dendrobium nobile* Lindl., an endangered medicinal orchid species. *Gene*. **529**:21–26.
- Börner, A., Plaschke, J., Korzun, V., Worland, A. 1996. The relationships between the dwarfing genes of wheat and rye. *Euphytica* **89**:69–75.
- Collard, B.C.Y., Mackill, D.J. 2009. Start codon targeted (SCoT) polymorphism: a simple, novel DNA marker technique for generating gene-targeted markers in plants. *Plant Mol. Biol. Rep.* **27**:86–93.
- Doyle, J.J., Doyle, J.L. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical. Bulletin* **19**:11–15.
- Dumolin-Lapegue, S., Demesure, B., Fineschi, S., Le Corre, V., Petit, R.J. 1997. Phylogeographic structure of white oaks throughout the European continent. *Genetics* **146**:1475–1487.
- Earl, D.A., von Holdt, B.M. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* **4**:359–361.

- Eticha, F., Bekele, E., Belay, G., Börner, A. 2005. Phenotypic diversity in durum wheat collected from Bale and Wello regions of Ethiopia. *Plant. Genet. Resour.* **3**:35–43.
- Etminan, A., Pour-Aboughadareh, A., Mohammadi, R., Ahmadi-Rad, A., Noori, A., Mahdavian, Z., Moradi, Z. 2016. Applicability of start codon targeted (SCoT) and inter-simple sequence repeat (ISSR) markers for genetic diversity analysis in durum wheat genotypes. *Biotechnol. Biotech. Eq.* **30**:1075–1081.
- Evanno, G., Regnaut, S., Goudet, J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* **14**:2611–2620.
- Gorji, A.M., Pocza, P., Polgar, Z., Taller, J. 2011. Efficiency of arbitrarily amplified dominant markers (SCoT, ISSR and RAPD) for diagnostic fingerprinting in tetraploid potato. *Am. J. Potato. Res.* **88**:226–237.
- Hamidi, H., Talebi, R., Keshavarz, F. 2014. Comparative efficiency of functional gene-based markers, start codon targeted polymorphism (SCoT) and conserved DNA-derived Polymorphism (CDDP) with ISSR markers for diagnostic fingerprinting in wheat (*Triticum aestivum* L.). *Cereal. Res. Commun.* **42**:558–567.
- Hamrick, J.L., Godt, M.J.W., Murawski, D.A., Loveless, M.D. 1991. Correlation between species traits and allozyme diversity: implications for conservation biology. In: Falk, D.A., Holsinger, K.E. (eds), *Genetic and Conservation of Rare Plants*. Oxford University Press. New York, USA. pp. 75–86.
- Hurd, E.A. 1971. Techniques for measuring plant drought stress. In: Larson, K.L., Rachter, J.D. (eds), *Drought Injury and Resistance in Crops*. CSSA Special Publication II, Crop Science Society of America. Cambridge University Press. Madison, WI, USA. pp. 77–88.
- Hutchenson, K. 1970. A test for comparing diversities based on the Shannon formula. *J. Theor. Biol.* **29**:151–154.
- IBPGR 1985. Revised Descriptor List for Wheat (*Triticum* spp.). International Board for Plant Genetic Resources. Rome, Italy.
- Moghaddam, M., Ehdai, B., Waines, J.G. 1997. Genetic variation and interrelationships of agronomic characters in landraces of bread wheat from southeastern Iran. *Euphytica* **95**:361–369.
- Moradkhani, H., Pour-Aboughadareh, A., Mehrabi, A.A., Etminan, A. 2012. Evaluation of genetic relationships of *Triticum-Aegilops* species possessing D genome in different ploidy levels using microsatellites. *Int. J. Agri. Crop. Sci.* **23**:1746–1751.
- Moradkhani, H., Mehrabi, A.A., Etminan, A., Pour-Aboughadareh, A. 2015. Molecular diversity and phylogeny of *Triticum-Aegilops* species possessing D genome revealed by SSR and ISSR markers. *Plant. Breed. Seed. Sci.* **71**:82–95.
- Naghavi, M., Malaki, M., Alizadeh, H., Pirseiedi, M., Mardi, M. 2010. An assessment of genetic diversity in wild diploid wheat *Triticum boeoticum* from west of Iran using RAPD, AFLP and SSR markers. *J. Agr. Sci. Tech.* **11**:585–598.
- Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetic* **89**:583–590.
- Pathak, N.N., Nema, D.P. 1985. Genetic advance in landrace of wheat. *Indian. J. Agri. Sci.* **55**:478–479.
- Peakall, R., Smouse, P.E. 2006. GENALEX 6: genetic analysis in excel. Population genetic software for teaching and research. *Mol. Ecol. Notes.* **6**:288–295.
- Perrier, X., Flori, A., Bonnot, F. 2003. Data analysis methods. In: Hamon, P., Seguin, M., Perrier, X., Glaszmann, J.C. (eds), *Genetic Diversity of Cultivated Tropical Plants*. London, Plymouth, UK. pp. 43–76.
- Pocza, P., Varga, I., Laos, M., Cseh, A., Bell, N., Valkonen, J.P. Hyvonen, J. 2013. Advances in plant gene-targeted and functional markers: a review. *Plant. Methods* **9**:1–31.
- Pour-Aboughadareh, A., Mohmoudi, M., Ahmadi, J., Moghaddam, M., Mehrabi, A.A., Alavikia, S.S. 2017. Agro-morphological and molecular variability in *Triticum boeoticum* accessions from Zagros Mountains, Iran. *Genet. Resour. Crop. Evol.* **64**:545–556.
- Powell, W., Morgante, M., Andre, C., Mm, Hanafey, Vogel, J., Tingey, S., Rafalski, A. 1996. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Mol. Breed.* **2**:225–238.
- Pritchard, J.K., Stephens, M., Donnelly, P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155**:945–959.
- Rajesh, M.K., Sabana, A.A., Rachana, K.E., Rahman, S., Jerard, B.A., Karun, A. 2015. Genetic relationship and diversity among coconut (*Cocos nucifera* L.) accessions revealed through SCoT analysis. *3 Bitech.* **5**:999–1006.

- Rawashdeh, N.K., Haddad, N.I., Al-Ajlouni, M.M., Turk, M.A. 2007. Phenotypic diversity of durum wheat (*Triticum durum* Desf.) from Jordan. *Genet. Resour. Crop. Evol.* **54**:129–138.
- Rholf, F. 2000. Numerical Taxonomy and Multivariate Analysis System: version 2.10. Exeter Software, Setauket. New York, USA.
- Salimi, A., Ebrahimzadeh, H., Taeb, M. 2005. Description of Iranian diploid wheat resources. *Genet. Resour. Crop. Evol.* **52**:351–361.
- van Slageren, M.W. 1994. Wild wheats: a monograph of *Aegilops* L. and *Amblyopyrum* (Jaub. & Spach) Eig (*Poaceae*). Wageningen Agricultural University Papers. Wageningen, The Netherlands.
- Wolfinger, R.D., Federer, W.T., Cordero-Brana, O. 1997. Recovering information in augmented designs, using SAS PROC GLM and PROC MIXED. *Agron. J.* **89**:856–859.
- Wright, S. 1951. The genetical structure of populations. *Ann. Eugen.* **15**:323–354.
- Xiong, F., Zhong, R., Han, Z., Jiang, J., He, L., Zhuang, W., Tang, R. 2011. Start codon targeted polymorphism for evaluation of functional genetic variation and relationships in cultivated peanut (*Arachis hypogaea* L.) varieties. *Mol. Biol. Rep.* **38**:3487–3494.
- Yeh, F., Yang, R., Boyle, T. 1997. POPGENE, the User Friendly Shareware for Population Genetic Analysis. Molecular Biology and Biotechnology Center. University of Alberta. Edmonton, Canada.
- Zhang, Z., Gao, J., Kong, D., Wang, A., Tang, S., Li, Y., Pang, X. 2015. Assessing genetic diversity in *Ziziphus jujuba* ‘Jinsixi-aozao’ using morphological and microsatellite (SSR) markers. *Biochem. Syst. Ecol.* **61**:196–202.

### Electronic Supplementary Material (ESM)

Electronic Supplementary Material (ESM) associated with this article can be found at the website of CRC at <http://www.akademaii.com/content/120427/>

#### Electronic Supplementary

Electronic Supplementary *Table S1*. Characteristics of sampling locations of Iranian *Triticum* accessions in this study

Electronic Supplementary *Table S2*. Mean, maximum, minimum, standard deviation (SD) and coefficient of the variation (CV) values for each traits evaluated in *Triticum* species

Electronic Supplementary *Table S3*. The first six principle components for 70 accessions of *Triticum*

Electronic Supplementary *Table S4*. Analysis of molecular variance (AMOVA) based on SCoT markers for four *Triticum* species

Electronic Supplementary *Table S5*. Genetic diversity estimates based on 15 SCoT primers in different *Triticum* species

Electronic Supplementary *Figure S1*. Two-dimensional plot derived from the PCoA of 70 accessions of *Triticum* based on SCoT data

Electronic Supplementary *Figure S2*. Graphical representation of genetic structure of four populations of Iranian *Triticum* inferred from an analysis of SCoT markers using the Bayesian clustering software STRUCTURE ver. 2.3. (A) Evanno test results showing estimated delta k. (B) Bar plots indicating representation of ancestry membership coefficients of all accessions in the four inferred genetic clusters: Red cluster (*T. boeoticum*), green cluster (*T. urartu*), blue cluster (*T. durum*) and yellow cluster (*T. aestivum*). They-axis shows the proportion membership into various clusters (for explanation of accessions codes, see Table S1). (C) The average membership fractions of the accessions from the four species of *Triticum* in the four inferred genetic clusters