Genetic diversity and relationships among traits in potato genotypes using agronomic traits and molecular marker (SSR)

Hossein Afshari[1*], Ehsan Barzin[1], Ghanbar laei[2], Morteza Noryan[3]

1. Department of Horticulture, Damghan Branch, Islamic Azad University, Damghan, Iran
2. Department of Agriculture, Damghan Branch Islamic Azad University, Damghan, Iran
3. Department of Agriculture, Azadshahr Branch, Islamic Azad University, Azadshahr, Iran

Abstract

The molecular marker (SSR) has been used to investigate the markers associated with the agronomic traits including days to 50% flowering, tube ring time, days to maturity, plant height, the number of main stems per plant, the number of tubers per plant, dry matter content, main stem diameter, a single tuber weight, average single tuber weight, and the total yield in potato genotypes. Ten primers used in 16 potato genotypes were investigated and principal component analysis and cluster analysis were used to determine genetic distance and genotype classification. A significant difference was found between all cultivars, indicating high genetic diversity for the traits under study, except for the diameter of stem trait. Yield showed a significant positive correlation with the number of tubers per plant, weight of tubers per plant, single tuber weight, and the number of main stems per plant. Special values resulted from components of 1-5 were higher than 1 in principal component analysis and totally justifying 76.70% of all variables variances. Molecular diversity of cultivars was measured using parameters of polymorphic information, the number of effective alleles, and Shannon index and the average of 39.03 SSR bands was achieved among which 9.41 bands were polymorphic. The average number of polymorphic bands varied for each primer from 1.56 to 2.12. Maximum and minimum polymorphic bands belonged to the HVM70 primer with 2.12 and the Bmacoo40 primer with 1.56, respectively. Also, the mean maximum band (4.4) and minimum band (3.1) belonged to Agra and Deyta cultivars, respectively. The morphologic data did not conform to the molecular data indicating that SSR marker had no genetic relationship with positions controlling the measured morphologic traits. Since the SSR markers were located in non-coding area of genome, no relationship between classifications of molecular markers and morphologic data was unexpected.

Keywords: primer; potato; Shannon Index; SSR markers; polymorphic bands

Afshari, H., E. Barzin, Gh. Laei and **M. Noryan.** 2017. 'Genetic diversity and relationships among traits in potato genotypes using agronomic traits and molecular marker (SSR)'. Iranian Journal of Plant Physiology 7 (3), 2095-2103.

Introduction

Today, in addition to increasing yield, seed production is particularly important and it has been considered as one of the main goals to introduce new cultivars because high yielding cultivars should have a desirable potential of seeding. Therefore, for identifying high yielding cultivars, it is necessary to identify traits that have

*Corresponding author
E-mail address: h_afshari@ymail.com
Received: December, 2016
Accepted: May ,2017

a significant relationship with yield so that the desirable genes in the improved cultivars are accumulated through selection (Bradshaw and Bonier bale, 2010). Legesse et al. (2007) stated that the molecular markers can be used as a useful tool in predicting hybrid performance when genes controlling the important traits saturated on the linkage maps are selected and the markers have high correlation with the QTL related with yield.

Simple Sequence Repeat microsatellite (SSR) is a repeated sequence of DNA which generally consists of 1-6 units of nucleotides. Because of the high capability of SSR in showing polymorphism, co-dominant inheritance, and relative frequency in the genome of these markers for genetic analysis, classification studies, investigating genetic diversity, composition of genetic maps, and selection of suitable parents before crossing different species are appropriate (Mohesni Fard et al., 2011; Dreisigacker et al., 2005 and Salimi et al., 2016).

In recent decades, many studies have suggested a positive relationship between genetic distances estimated by molecular markers and heterosis. For example, Mohsenifard et al. (2011) used 10 pairs of SSR molecular marker to investigate the relationship between lines, the genetic distance, and the selection of suitable parents for the production of hybrid varieties of tomato in 16 lines. The results showed that SSR molecular marker was efficient for the diagnosis of polymorphism and the genetic distance between the lines of tomato and it can be used as an accurate tool in the analysis of genetic relationships among the samples. However, the study found no significant relationship between the genetic distance estimated by SSR molecular marker and heterosis of 63 hybrids evaluated for yield, fruit size, fruits number, and fruit shelf life. Many studies have been conducted on a wide range of agricultural products to predict the heterosis and the efficiency of hybrids' performance by using genetic diversity based on the various molecular markers. Studies carried out in this field have showed conflicting results. For example, there was a positive and significant correlation between genetic distance based on the molecular markers and heterosis hybrids in studies conducted on potato (Bahar *et al.,* 2007), tomato (Mirshamsi Kakhaki, 2002), wheat (Maghrebi et al., 2005) on wheat (Rural et al., 2005), barley (Triticale Varshney et al., 2009) and bread wheat (Yau et al., 2005). On the other hand, there was no significant correlation between genetic distance based on the molecular markers and heterosis on offspring in another set of studies on fava beans (Yau et al., 2005), wheat (Dreisi Gacker et al., 2005; Maghraby et al., 2005), and corn (Lex et al., 2007).

The present study investigated the genetic diversity of yield in several accessions of genotypes of potatoes using agronomic traits and molecular marker (SSR). The study also evaluated genotypes of potatoes to find high yielding varieties and the relationship between yield and other morphological traits in potato plants and to determine the genetic diversity pattern and classification of genotypes based on morphological traits and yield and also to calculate the genetic distance between genotypes using SSR molecular markers (Górall et al., 2005; Galaev et al., 2006)

Materials and Methods

The study was conducted in 2015 in Dezful located between 48 degrees and 20 minutes to 48 degrees, 31 minutes to east longitude and 32 degrees and 75 minutes to north latitude with the land area of 4762 square kilometers. Sixteen genotypes of potato were obtained from Department of Natural Resources and the Gene Bank of the Research Institute (Table 1). A factorial experiment was carried out in a randomized complete block design (RCBD). The potato tubers were planted after tillage in mid-May. Each plot consisted of two 6-meter lines located on two stacks and their distance with plants was 25 cm. Fertilizers were applied based on the soil test

Table 1
Potato genotypes used in the study

ROW	GENOTYPE	ROW	GENOTYPE
1	Picasso	9	397007
2	Marfona	10	397045
3	Data	11	397015
4	Casmus	12	396140
5	Arinda	13	397079
6	Agria	14	Sante
7	394600	15	Lady rosetta
8	396757	16	Mondial

results. The irrigation repeated every 7 days. During growth time, the traits for days to 50% flowering, tube ring time, days to maturity, and plant height were measured. After plantation and plant establishment, random sampling was carried out at five stages. Sampling was done after 50 days of plantation and it was repeated every 15 days. The last sampling was done along with the harvesting of plants. After harvesting, traits for the number of main stems per plant, the average number of tubers per plant, single tuber weight, mean tuber weight, and total yield per plant were measured and recorded. The DNA was extracted from young leaves at six-leaf stage about 5 weeks after sowing by CTAB, according to the Saghaei et al (1984). Quality and quantity of the extracted DNA using agarose 1% gel was determined by spectrophotometer. PCR program was set on Touch Down for the primers with junction temperature at 55° C and 60° C. To get started, the samples were run first with 2% gel and then, with the vertical gel in order to confirm them. Afterwards, the gels were colored using silver nitrate coloring method. Polymorphism information was investigated using the following equation:

$$\mathrm{PIC} = 1 - \sum_{i=1}^{n} P_i^2$$

where P_i is the frequency of alleles for each population and n is the number of effective alleles.

Shannon index was also calculated using the following equation:s

$$H' = \sum P_i \ln P_i \quad \text{and} \quad P_i = \frac{n_i}{N}$$

In this equation P_i is the frequency of alleles for each population, n_i is the number of effective alleles indicating the genetic diversity in the population, and is used to compare the populations.

The effective number of alleles was calculated by using the following equation:

$$\mathrm{Ne} = \frac{1}{L}\sum_{1}^{L} = 1\left(\sum_{a=1}^{n_1} P\text{^}2\right)^{-1}$$

In this equation, L is number of loci (markers), n_1 is the number of alleles per locus 1 and P is the relative frequency of a in locus 1.

The similarity matrix was provided using DIS method in order to do cluster analysis based on the molecular data. Then, a dendrogram was drawn by the algorithm of the neighbor joining. The molecular traits were investigated using molecular marker SSR, according to Roder et al. (1998). Cluster analysis on the genetic distance was done based on the molecular markers by Un-weighted Pair-Group Method with Arithmetic Averages (UPGMA). In addition to the SSR molecular marker, information related to morphological traits of genotypes was used to classify the 16 genotypes of potatoes. Classification was done using the markers and morphological traits through STATISTICA software based on Euclidean distance.

Results

Principal component analysis

The aim of principal component analysis (PCA) is data reduction. In this method, the relationships between the traits can be found through investigating the correlation between the variables. In PCA, the correlation between a large number of variables is expressed by several independent factors and the role of each trait is determined in the existing diversity. In addition, principal components analysis is used for classification of cultivars and in fact, it is a supplementary method for cluster analysis (Singh and Chaudhary, 1985). Principal component analysis is usually done before the cluster analysis to determine the relative importance of variables that are involved in the classification of genotypes in cluster analysis (Jackson, 1991).

Given the diversity of traits, 10 primers were used on 16 genotypes in principal components analysis (Table 2). Parameters of principal component analysis are presented for the components 1 to 5 for each of the three cases in Table (3) where their parameters include eigenvalues, percentage of the justified variance, and coefficients of specific vectors. Eigenvalues resulted from the components 1 to 5 were more than one, and their percentages were 24, 12, 22,

11 and, 8 respectively. They were justified 76% from the variables' total variance. The relative values of the coefficients of eigenvalues in the first component showed that tubering time, days to 50% flowering, yield, and dry matter content were the most important traits for classifying the genotypes in the cluster analysis. In the second component, days to 50% flowering, plant height, and number of tubers per plant and in the third component, the average weight of tubers per plant had higher coefficients of specific vectors. Similarly, in the fourth component, the main stem diameter and the number of main stems per plant, and in the components 1 to 5 the phenology component, yield component, component of height, and single tuber weight parameters were named, respectively (Table 3).

Cluster analysis

In plant breeding, efficiency of selection depends on the diversity, genetic recombination, and heterosis. There are several reports that the likelihood of heterosis will increase in cross programs by increasing genetic distance between the genotypes of specie (Humphreys, 1991). Classifying the genotypes based on the genetic distance is effective in breeding plan when several traits are evaluated simultaneously. Therefore, cluster analysis was performed in order to determine the genetic diversity pattern, classify the genotypes, and determine the genetic distance between them.

Table 2
Primers used in the study

ROW	Name
1	HvWaxy4a
2	Bmac0316
3	Bmac0040
4	EBmac0415
5	HVHVA1
6	Bmag0603
7	Bmag0013
8	GMS001
9	HVM70
10	HVM40

Cluster analysis was used for 10 traits in 16 genotypes. In Table (4), the number of clusters and their genotypes are listed. The cluster analysis of genotypes was used for 10 primers and the genotypes were divided into 7 groups by cutting dendrogram

According to the results, clusters 1 and 6 had better yield and genotypes in clusters 1 and 6 are recommended for cultivation. Distribution of 16 genotypes of potatoes based on two main components is showed in Figure (I). In this figure, the first component in terms of tubering time, days to 50% flowering, yield, and dry matter played an important role in differentiating the groups (Table 4)

Table 3
Eigenvalues, percentage of variance, and coefficients of specific vectors for each studied trait in the principal component analysis on means of data

Trait	PCA1	PCA2	PCA3	PCA4	PCA5
Tubering time	.34	-.33	.33	.0	-.08
Days to 50% flowering	.39	.33	.17	.13	.13
Plant Height	-.10	.38	.26	-.28	.22
Main stem diameter	-.28	.21	.05	.46	.10
Weight of single tuber	.10	-.01	.11	.11	.63
Average of single tuber in plant	-.029	-.31	.40	.10	-.10
Yield	.40	.21	.24	.25	-.20
Number of tuber per plant	-.17	.44	.34	.13	.22
Number of stems in plant	-.19	.21	.06	.60	-.15
Days to maturity	-.08	.18	.15	.00	-.61
Dry matter content (%)	.38	-.07	-.14	.36	-.03
Eigenvalues	3.41	2.42	1.97	1.63	1.29
Total Variance	24.39	17.31	14.08	11.68	9.24
Cumulative Variance (%)	24.39	41.70	55.78	67.45	76.70

More valuable numbers in main components are underlined in the table.

Table 4
Potato genotypes in each clusters

1	2	3	4	5	6	7
Agria	397007	396140	394600	Sante	Mondial	Marfona
Casmus	397045	397079	Picaso	Lady		Data
Armida				397015		
				396157		

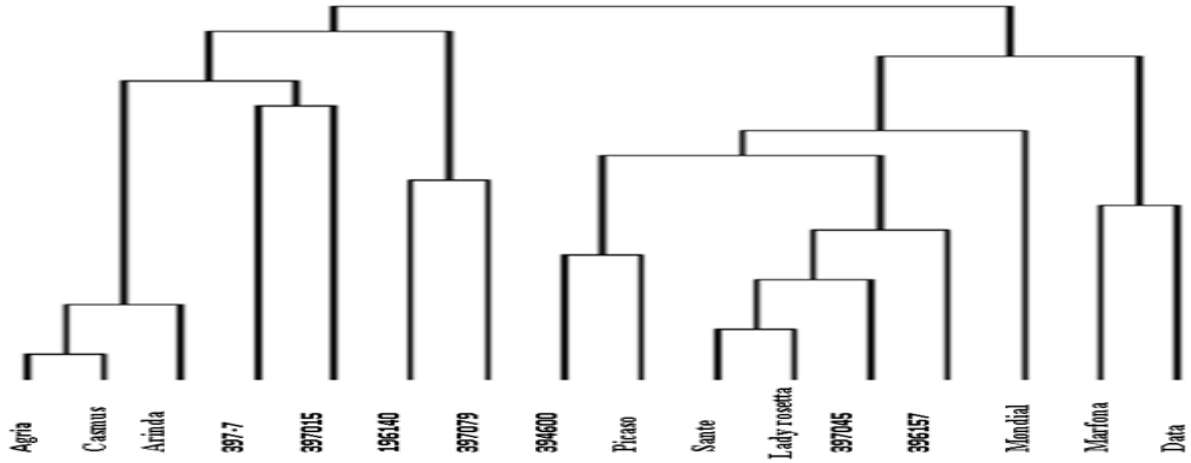

Fig. I. Dendrogram resulted from cluster analysis on 16 genotypes of potatoes

In the second component, plant height, days to 50% flowering (with positive sign) and the number of tubers per plant (with a negative sign) played a role in differentiating the groups, so that groups with the higher number of main stem, seed yield, and number of tubers per plant were less scattered at the bottom of the chart. Accordingly, cluster 1 on the highest part of the chart had a high yield potential, and if the purpose of breeder is increasing the yield, it is advised to use the genotypes in cluster 1.

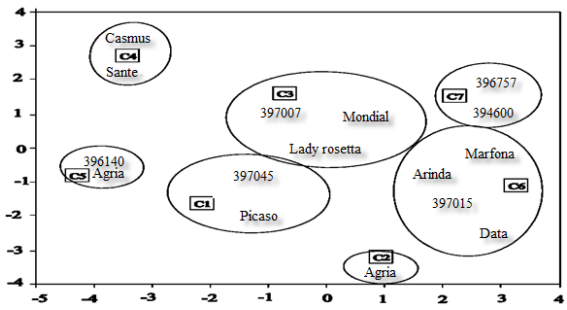

Fig. II. Scatter diagram of 16 potatoes varieties based on the two main components

Table 5
Primers' polymorphism information and the number of bands in 16 varieties of potato

Mean	Number of band	Polymorphic bands	Percentage of polymorphic	Polymorphic coefficient(PIC)
Gms001	4.12	2.00	47.49	.86
BMac0316	4.00	2.06	51.24	.82
Bmac0040	3.93	1.56	41.45	.81
EBmac0415	4.06	1.87	46.97	.84
HVHVA1	4.00	2.06	51.56	.83
Bmag0603	3.81	1.87	48.95	.81
Bmag0013	3.68	2.00	55.51	.76
HVwaxy4a	4.00	2.00	50.51	.82
HVM70	3.93	2.12	54.37	.82
HVM40	3.50	1.87	54.99	.71
Total Means	39.03	19.41	503.04	8.08

Demirel et al. (2017) conducted factor analysis using the traits of the tubers. In this study, the first factor showed a negative structural relationship between the interior and the membrane parts and the second factor indicated a positive relationship between the width indicators and length of the tuber.

Cardi et al. (2002) also identified 10 first main components through justifying %98.9 of total variance by using PCA on tetraploid cultivars of potato. These components often highly correlate with the number, size, and shape of the leaf, flowering, flower shape, and the number of stems.

Molecular analysis and means comparison

Molecular analysis of 16 potato genotypes using SSR markers were scored on the average 39.03 bars among which 19.41 bars were polymorphic.

The maximum mean number of polymorphic bands was related to the primers BMac0316 and HVHVA1 with 2.06 bands and the minimum mean number of polymorphic bands was related to the primer Bmacoo40 with 1.56 bands. The lowest percentage of polymorphism (41.45) was recorded for the primer Bmac0040 and the highest percentage of polymorphism (55.51) was recorded for the primer Bmag0013. The mean percentage of polymorphic of the primers was 503.04. Moreover, 61.11% of total bands was equal to 397045 for all polymorphic primers, and this figure had the most variety. Also, 41.46% of polymorphism for all primers was equal to 396140 which showing the lowest diversity. The mean of maximum bands belonged to Agria with 4.4% bands while the minimum belonged to the data with 1.3 bands. Polymorphism coefficient varied from 0.51 to 1.4. With respect to the polymorphism, the number of bands, and polymorphic coefficients of the most effective primer, Bmag0013 created more diversity (Table 5).

According to the results, the maximum polymorphism percentage belonged to 397045 with 61.11% and the minimum polymorphism percentage belonged to 396140 with 41.46%. Also, mean of maximum bands belonged to Agria

Table 6
Number of the observed alleles

Variety	Numbers of Observed Allele
Picaso	2
Marfona	3
Data	3
Casmus	2
Armida	4
Agria	2
394600	3
396157	2
397007	2
397045	2
397015	2
396140	3
397079	2
Sante	3
Lady rosetta	3
Mondial	2
Total	40
Average	2.5

Table 7
Number of the effective alleles

Variety	Numbers of Effective Allele
Picaso	2.25
Marfona	3.90
Data	3.90
Casmus	2.25
Armida	4.50
Agria	2.25
394600	3.90
396157	2.25
397007	2.25
397045	2.25
397015	2.25
396140	3.90
397079	2.25
Sante	3.90
Lady rosetta	3.90
Mondial	2.25
Total	40.15
Average	3.00

Table 8
The mean observed and effective alleles and Shanon Index in 16 genotypes of potato

	Total Cultivars
Numbers of Observed Allele	2.5
Numbers of Effective Allele	3.00
Shannon Index	.76

Table 9
Genotypes of potato in clusters

Cluster number		
1	2	3
Armida	396157	Sante
394600	397045	Lady rosetta
397007	Picaso	Marfuna
397015		Data
396140		Casmus
397079		Agria
		Mondial

with 4.4 while mean of minimum bands belonged to the data with 3.1.

The number of observed and effective Alleles and Shannon Index

Some parameters related to genetic diversity such as the number of observed alleles, the number of effective alleles, and Shannon index (H') were calculated for the total genotypes (See Tables 6, 7, and 8). The average observed alleles for all genotypes was 2.5. The average number of effective alleles for all genotypes was 3.0 and Shannon index for all genotypes was .0.76 (Table 8). As the table suggests, the number of effective and observed alleles in the Armida was more than other varieties and this implied the level of polymorphism for total Primers. Due to the higher coefficient number in 397045, polymorphisms of this number was higher for all primers

Cluster analysis based on the molecular data

Based on the classifications of genotypes and estimation of the average of traits for each cluster and the total mean deviation, suitable parents can be selected in order to create variation in breeding plans. Since the genotypes available in each cluster have more genetic similarity than other clusters, classification based on the molecular data can be used to create more diversity and conducting targeted crosses. This is because the effect of the environment is eliminated in classification and a more accurate classifying of genotypes can be achieved.

Cultivars were classified into 3 clusters in cluster analysis based on the molecular data by cutting dendrogram (See Table 9). These include:

Genotypes in cluster 1: Armida, 394600, 397007, 397015, 396140, 397079,

genotypes in cluster 2: 396,157, 397,045, Picasso, and

genotypes in cluster 3: Sante, Lady Rosetta, Marfuna, Data, Cosmus, Agria, Mondial.

The first cluster consisted of Armida, 394600, 397007, 397015, 396140, and 397079,

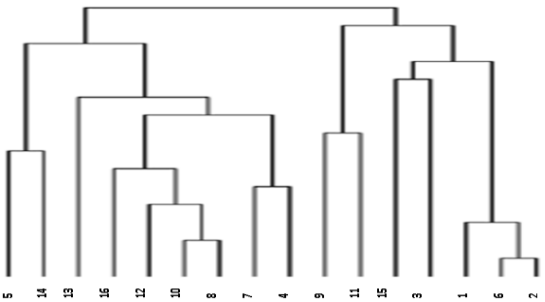

Fig. III. Cluster analysis of 16 varieties of potatoes based on the molecular data

which were in the same cluster in cluster analysis based on morphological traits. In the second cluster, cultivars of 396157, 397045, and also Picasso were in the same cluster in cluster analysis based on morphological traits. But the third group was different at the molecular and DNA level of genotypes or in insufficient number of markers. Also, the difference between clustering based on the morphological data can be due to the effects of environment on phenotype.

Discussion

One of the main objectives of plant breeding is increasing the yield per unit area. Results of mean comparison in potato genotypes based on molecular data showed that the average maximum band was produced in Agria genotype, and the highest average percentage of polymorphism was related to genotype 397045. Also, with respect to polymorphism, the number of bands and the polymorphic coefficient value of the most effective primer was Gmsoo1 which showed the highest diversity. There were positive and significant correlations between yield and tuber weight per plants, tuber wet weight per plant, and single tuber weight, and this implies that the yield will increase when fertile stems are increased. In principal component analysis on the mean data of the environment, eigenvalues of the components 1 to 5 were more than one, and they generally justified 76.70% of total variance of the variables. According to the results of the second component, plant height, days to flowering and the number of tubers per plant and also results of the third component, average of tuber wet weight per plant had higher specific vector coefficients.

Coefficients of various specific vectors in independent variables showed that it is possible to

improve the yield of potatoes by selecting different combinations of these qualitative and quantitative traits. The genotypes were divided into 7 different groups in cluster analysis based on the morphological traits. Varieties in cluster 6 were recommended for cultivation. Armida was known as the desirable cultivar which had a good yield. The genotypes were also divided into three groups in cluster analysis based on molecular data. Classification of morphological data was not consistent with the molecular data. This showed that the SSR markers used in this study did not have appropriate genetic relation and linkage with controlling places of the morphological traits studied. But, since the SSR markers are mainly located in non-coding regions of the genome, disaffiliation between classifications of molecular data and morphological data were predictable. The most important trait in breeding plans is yield and the traits related to the yield are also important. As it was observed, there was positive and significant correlation between seed yield and the number of tubers per plant, tuber weight per plant, single tuber weight, and the number of main stems per plant. In principal component analysis, tubering time, days to 50% flowering, and dry matter content were classified in the same component with yield. It is necessary to consider these traits to improve the yield in breeding plans.

References

Bahar M, H. Mohammadi and **S. Ghobadi**. 2007. 'Evaluation of potato genotypes genetic diversity by SSR markers. *J Sci Technol Agric Nat Resour*, 4:271-279.

Bradshaw J, E and **M. Bonier bale.** 2010. 'Potatoes'. In: Bradshaw, J.E. (ed.). Root and Tuber Crops. Handbook of Plant Breeding Vol. 7, Springer, New York.

Cardi, T., M. Mazzei and **L. Frusciante**. 2002. 'Field variation in a tetraploid progeny derived by selfing a *Solanum commersonii*(+) *S. tuberosum* somatic hybrid: A multivariate analysis'. *Euphytica*, 124(1): 111-119.

Demirel, U., İ. Tındaş, C. Yavuz, F. Baloch and **M. Çalışkan.** 2017. 'Assessing genetic diversity of potato genotypes using inter-PBS retrotransposon marker system'. *Plant Genetic Resources*, 1-9.

Dreisigacker, S., A.E. Melchinger, P. Zhang, K. Ammar, C. Flachenecker, D. Hoisington and **M.L. Warburton.** 2005. 'Hybrid performance and heterosis in spring bread wheat and their relations to SSR-based genetic distances and coefficients of parentage'. *Euphytica*, 144: 51-59.

Galaev A. V., L. T. babayants and **Yu. M. Sivolap**. 2006. 'DNA-markers for resistance to common bunt transferred from *Aegilops*., *cylindrical* Host to hexaploid wheat, *Czech* J. Genet. Plant Breed 42:65-62.

Górall, H., M. Tyrka and **L. Spiss**. 2005. 'Assessing genetic variation to predict the breeding value of winter triticale cultivars and lines'. *Theoretical and Applied Genetics*, 46:125-131.

Humphreys, M.O. and **C.F. Eagles.**1987. 'Assessment of perennial ryegrass (*Lolium perenne* L.) for breeding. I. freezing tolerance. *Euphytica*, 38:75-84.

Jackson, J. E. 1991. 'A user's Guide to Principal Components'. Willy, New York.

Legesse, B.W., A.A. Myburg, K.V. Pixley, S. Twumasi- Afriyie and **A.M. Botha**. 2007. 'Relationship between hybrid performance and AFLP based genetic distance in highland maize inbred lines'. *Euphytica*, 162:313-323.

Maghraby, M.A.L., M.E. Moussa, N.S. Hana and **H.A. Agrama.** 2005. 'Combining ability under drought stress relative to SSR diversity in common wheat'. *Euphytica*, 141: 301–308.

Mir Shamsi Kakhki, A. 2003. 'Evaluating the potential of RAPD markers in estimating heterosis and the combination ability of line and cultivars of tomatoes to produce hybrid seed'. M.A. Thesis, Mashhad: Ferdowsi University, Faculty of Agriculture.

MohesniFard, A., M. Farsi, S. H. Nemati and **Kh. MalekZade**. 2011. 'Assessment of tomatoes genetic diversity using molecular marker (SSR) and its correlation with the heterosis'. *Iran: Journal of horticultural sciences*, 42: 185-192.

Roder, M. S., V. Korzun, K. Wendehake, J . Plaschke, M. H. Tixier, P. Leroy and **M. W. Genal.**1998. 'A microsatellite map of wheat'. *Genetice*, 149, 2007-2023

Salimi. H., B. Masoud, A. Mirlohi and **M. Talebi** .2016. 'Assessment of the genetic diversity

among potato cultivars from different geographical areas using the genomic and EST microsatellites'. *Iran J Biotech*. 14(4): e1280.

Singh R. K. and **B. D. Chaudhary.** 1985. 'Biometrical methods in quantitative genetics analysis'. Publisher Kalyani Publishers, 318 pages.

Varshney, R. K., M. Baum, P. Guo, S. Grando, S., Ceccarelli and **A. Graner**. 2009. 'Features of SNP and SSR diversity in a set of ICARDA barley germplasm collection'. *Mol. Breed*. 26: 229-242.

Yau, S.k., G.O. Ferrara and **P. Srivastava**. 2005. 'Classifiction of diverse bread wheat growing environments based on differential yield response'. *Crop Sci*. 31: 571-576.