A study on genetic diversity in lentil genotypes using seeds morphologic and protein traits

Parisa Aghili*, Ali Akbar Imani and Yousef Alaei

Department of Agronomy and Plant Breeding Ardabil branch, Islamic Azad University, Ardabil, Iran

Corresponding author: Parisa Aghili. Department of Agronomy and Plant Breeding Ardabil branch, Islamic Azad University, Ardabil, Iran. Email: parisaaghili@yahoo.com

Abstract: The following research tries to study the relation and correlation between grain yield and other quantitative traits in lentil using 29 lentil genotypes (including 26 foreign genotypes and 3 control genotypes). The research was conducted in Ardabil Agriculture and Natural Resources Research through augmented method in randomized complete block design in three replications, during 2011. During the agricultural season, certain traits such as green percentage, days to flowering, number of hooks, hook size and grain yield were measured. Subsequent to the variance analysis, data related to the control cultivars, and also estimation of blocks effects and amending each studied treatment on the studied traits, the relation between evaluated traits and grain yield were studied. Results suggested that there is a positive significance relation between the green percentage, hook size, plant height, 100 pods weight, 100 seeds weight, biomass and number of filled pods on the one hand and the grain yield on the other. Step-by-step multiple Regression results indicated that among the studied traits, biomass and number of secondary branches explain more than 84% of the grain yield changes so that, the increase in biomass and decrease in number of secondary branches, increase the yield. Cluster analysis divided studied genotypes into three groups in which, the first group with genotype numbers of 1, 5, 6, 9, 10, 11, 14, 15 and 21 was the best group. According to the protein data, the highest number of protein band (22) were observed in genotype numbers of 8, 21 and control genotype number of 27 while the lowest number of protein band (16) were observed in genotype numbers of 19 and 20, so that bands numbers of 2, 3, 5, 7, 8, 13, 14, 15 and 16 with respective molecule weight of 118.35, 112.71, 99.77, 86.17, 80.09, 44.58, 42.46, 40.43 and 38.51 KD a were diagnosed as polymorphism bands. According to the protein data, genotypes were divided into three groups in which the third group with 12 genotypes of 1, 2, 3, 4, 5, 6, 11, 12, 13, 15, 20 and 22 had a higher value as the delayed, high yielding and long-legged genotypes along with most of studied traits. The farthest distances from protein bands were related to the genotypes numbers of 23 with 14, 17, 18 and 19. Results suggested that grouping based on morphologic data was to 35% consistent with protein data.

[Parisa Aghili*, Ali Akbar Imani and Yousef Alaei. **A study on genetic diversity in lentil genotypes using seeds morphologic and protein traits.** *Life Sci J* 2012;9(4):4099-4106]. (ISSN: 1097-8135). http://www.lifesciencesite.com. 609

Keywords: genetic diversity, morphological traits, protein, lentil, electrophoresis

Introduction

Morphological indicators indicate the variety in shape or yield in plants. Emergence of awn, pigments, reaction to hormones, herbicides and diseases are among such indicators. However, phenotypic assessments have limited application due to the environment effects on gene expression, dominance and epistatic effects, presence of pleiotropy, changes in gene penetration, dependence on the tissue and developmental stages, assessment tests being time consuming and the limited genetic information obtained (Musavizade, 2006). Protein indicators represent the variation in protein products of genes. Isozyme and endosperm protein compounds are of this type. There are some biochemical methods presented based on electrophoresis of seed proteins and enzymes whose usefulness have been proved in the analysis of genetic diversity. Using various alleles of one or multi-locus forms, these methods identify the differences between seed storage proteins or coded enzymes. Using biochemical methods could omit the environmental effects. However, its usefulness is limited due to its inability to detect low levels of diversity, limited genome coverage, non-random distribution and its limitations in number (Bozorgi, 1994). In most cases, seeds are considered as the sources for protein, for they represent a certain stage in a plant lifetime. For instance, varieties related to a leaf growth could limit their protein pattern for taxonomic purposes. In addition, seeds are great protein sources and obtain enough protein for electrophoresis. The main reason to use seed stored proteins electrophoresis patterns in categorization is due to proteins being relatively direct products of genes. Hence, it is believed that these patterns could represent criteria in genetic similarities and differences among comparing plants. Using seeds protein patterns in systematic studies is based on this

assumption that proteins of various individuals, various populations and various species are similar if they maintain a similar move in a gel and they produce bands with almost the same width and intensity after staining. Each band is studied as a separate trait and it is assumed that these traits are the relatively direct products of genes. The main method for assessing protein similarities among populations and taxa is to use a similarity criterion (Rahiminejad, 1999). Simple counting of the ratio in which "a" indicates the number of common bands and "b" indicates the total of bands found in two populations or taxa, is the common method for showing the protein similarities. It should be mentioned that this method does not lead into genetic distance (Sahai and Rana, 1977). Seeds proteins electrophoresis is a suitable method for obtaining systematic quantitative information from macromolecules. Also, the pollen protein is used in few cases. Sometimes a mixture of protein essences related to two taxa is put in the stream to assess if bands separate or have a side by side migration. This method could provide the potentiality for a more pure side by side assessment, comparing to putting separate essences and side by side in a gene form (Rahiminejad, 1999). Seed protein profile could contain 20 or more single bands. Band patterns complexity could result in difficulties in interpreting information. Also, by increasing in the number of band and studied populations, the accuracy must be increased, as well. Considering the aforementioned difficulties in seeds protein profiles scores, the stability of mature seed storage protein stability is not affected by the seasonal, environmental and seed longevity fluctuations. Also, these profiles are unique to each species (Ladizinskey and Hymowit, 1979). The main objectives in this research include studying genetic affinity in some lentil genotypes, using seeds morphological traits and seeds protein storages.

Material and methods

In this experiment were used of lentil genotypes (including 26 foreign genotypes and 3 control genotypes). Seed samples were produced, of beans Research Center Agriculture and Natural Resources in Ardebil province east Cost of Iran.

Experimental procedure

We used Pilot project used augmented design as a randomized complete block design with three replications a split-plot design with three replications

Traits

Traits Average based on 10 plant competitors who were randomly selected and analyzed following measurements: green Percent, days to flowering time, number of hooks, hook size, grain yield per unit area, days to reach a plant height, Height, lowest pod, harvest index, number of filled pods per plant, empty pods per plant, seed number per 100 pods, number of primary branches per plant, number of secondary branches per plant, biomass and seed weight.

Protein Extraction

In this stage, 20 healthy and medium seeds from each genotype are selected and after separating lemma and palea, they were pounded between oil-paper. The pounded materials from each genotype were poured in an Eppendorf pipette and each sample specifications were recorded on each sample. 400 microliters of the extracted solution were added to each sample. (0.1 mililiter of solution for 8 mg of the sample) While gels were polymerized, protein extraction operation was done. After adding extracted solution on pounded samples, Eppendorf pipettes are immediately shaken by shaker so that the pipettes contents are fully mixed. During the two hours of protein extraction, the aforementioned operations were done 3 to 4 times until the protein extraction was fully done. After two hours, centrifuge was done for 10 minutes in 10,000 rpm at 4 °C. Solid matters were completely settled, after centrifuge. 200 microliters of supernatant was taken form the solution on Eppendorf pipettes and transferred to the new Eppendorf pipettes by preserving the genotypes traits (extracted protein was preserved at -20 °C)

Proteins Electrophoresis Part

In electrophori studies also, 29genotypes were studied. It is proved that seeds storage protein variety is used in SDS-PAGE for identifying various genotypes and the most common technique used for analyzing mixed protein is the SDS-PAGE method in which proteins are separated based on their sizes (Shehata, 2004).

Bands Identification:

Jaccard similarity coefficient is calculated from the following:

In which, "a" is the control bands in both species and "b" is the number of unique bands in the first species and "c" is the number of unique bands related to the second species.

Protein Bands Cluster Analysis

To conduct the analysis, bands zero and one matrix in NTSYS 2.02e was used. To determine the distance between genotypes simple matching similarity coefficient was used. And to merge the clusters UPGMA method was used.

Results and Discussion

Cluster Analysis Based on Morphological Traits

To study and categorize the studied cultivars, Ward method was used in cluster analysis based on assessed traits in 3groups. Specifications for each cluster are presented below:

- First group includes 9 genotypes (1, 5, 6, 9, 10, 11, 14, 15 and 21) which are high yielding and legged, and also, they obtain a high value in biomass, number of full pods, weight of 100 grains, pod lower height, green percentage, number of empty pods, number of primary and secondary branches among other clusters. (Table 1)
- Second group includes 17 genotypes (2, 3, 4, 7, 8, 12, 13, 16, 17, 18, 19, 20, 22, 23, 24, 25 and 26) which are late flowering and late crop, and also, they obtain a high value in the number of hooks, harvest index and primary and secondary branches among other clusters.
- Third group included the control genotypes (27, 28 and 29) which obtain lower values in all studied traits among studied genotypes.

Hence, it could be concluded that among the aforementioned groups, the first group with genotypes of 1, 5, 6, 9, 10, 11, 14, 15 and 21 is the best group.

Study of Seed Storage Protein Variety Using SDS-PAGE Electrophoresis

During this study, all stored proteins were extracted from seeds. The gel derived from total proteins electrophoresis was coded based on presence or absence of bands (protein pattern). Presence of band was presented by "1" and absence of bands was presented by "0" and a matrix was finally formed.

Protein bands map is presented in Figure 2. Number of bands presence according to the genotypes is presented in Table 2. Bands' molecular weight and their FRs are presented in Table 3. 23 bands were totally studied in this research, whose molecular weight had a change range between 17 to 127.5 KD and their RFs had a change range between 0.27 and 0.98.

The highest number of bands (22) was observed in genotypes of 8 and 21 and control genotype of 27 and the least number of bands (16) was observed in genotypes of 19, and 20. Band numbers of 12, 11, 10, 9, 6, 4, 1, and 17 to 23 with molecular weight of 49.15, 52.89, 58.31, 64.29, 92.72, 104.76, 127.35, 35.79, 33.29, 30.17, 26.06, 23.63, 20.92 and 18.97, respectively, were common between all genotypes. Other bands showed polymorphisms of presence or absence type. Bands number of 2, 3, 5, 7, 8, 13, 14, 15 and 16 showed polymorphisms with molecular weight of 118.38, 112.71, 99.77, 86.17, 80.09, 44.58, 42.46, 40.43, and 38.51, respectively.

Analysis of Seed Storage Proteins Cluster:

Various cluster analysis methods were reported for protein patterns data (Huff et al., 1993; Liu et al., 1994; Wu & Lin, 1994; Peakall et al., 1995; Huff, 1997). for choosing the categorization method, the Cophenetic coefficient was calculated using NTSYSc 2.02e software whose highest value was related to UPGMA method, based on Jaccard similarity matrix (r=0.79) (Jaccard, 1908). (Table 3)

According to the results, the 29 genotypes are divided into 3 groups, so that, the first group included 7 genotypes of G7, G8, G27, G21, G9, G10 and G23 which had a medium yield with medium traits.

The second group included 10 genotype's of G14, G28, G25, G26, G29, G24, G16, G17, G18 and G19 which were late flowering genotypes with highest harvesting index and weight of 100seeds. They were low in other studied traits.

The third group included 12 genotypes of G1, G22, G11, G13, G15, G12, G20, G2, G3, G4, G5 and G6 with highest number of members which were late crop, high yielding and legged. Also, they obtained high values in studied traits. (Table 4)

Comparing the results for categorizations derived from electrophoretic data cluster analysis and categorizations derived from morphologic data cluster analysis, it could be observed that around 10 genotypes are categorized in one group. In other words, categorizations based on morphological traits and protein bands had a consistency at 35%. Similarity coefficient between genotypes based on protein bands were calculated by Jaccard method:

$$\text{Jaccard Coefficient} = \frac{a}{a+b+c}$$

In which, "a" is the control bands in both species and "b" is the number of unique bands in the first species and "c" is the number of unique bands related to the second species (Moqaddam et al., 1994).

It should be mentioned that these coefficients vary between the range of 0.667 and 1. The higher the similarity coefficient between two genotypes, the more the similarity between two genotypes is higher based on protein bands and biochemistry. According to Table 5, the least similarity coefficient was between genotypes of 23 and 14, 17, 18 and 19 and 27 and 8, and 19 and 20 and 7 with 17 and 9 with 18 which show a great difference between genotypes on seed total proteins. To achieve the maximum HYTHROSIS in hybridizations, genotypes, which have the highest difference on protein bands electrophoretic patterns, are mixed.

Table 1: Average traits to distinguish clusters from a cluster analysis of genotypes

Group	Number genotype	Weight of 100grains	Bio mass	Secondary branches	Primary branches	Seeds in 100 pods	Number of empty pods	Number of full pods	Harvest index	Height lowest pod	Height	Days to reach	Yield	Hook sizes	Number of hooks	Flower development	Germination
1	9	04.7	02.1	00.4	00.4	52.265	60.7	80.51	44.60	65.13	82.42	11.2	85.2	00.2	45.2	62.75	27.90
2	17	64.6	72.7	23.4	23.4	47.259	70.10	89.43	85.70	38.13	64.42	14.2	83.1	00.2	47.2	52.78	44.75
3	3	11.4	77.4	00.2	00.2	83.125	68.3	11.23	08.28	90.6	98.20	06.1	34.1	00.1	33.1	33.75	50.92
Total	29	50.6	21.8	93.3	93.3	52.247	01.9	19.44	20.63	79.12	45.40	02.2	10.2	89.1	34.2	29.77	81.81

Table 2: Number of protein electrophoresis to separate bands of lentil genotypes

Genotype	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12	G13	G14	G15	G16	G17	G18	G19	G20	G21	G22	G23	G24	G25	G26	L1	L2	L3
Band	18	19	19	19	20	20	21	22	20	19	17	18	17	17	17	17	17	18	16	16	22	18	19	18	19	19	22	17	20

Table 3: bands observed in the electrophoresis of proteins - the molecular weight and relative mobility

Band	Number of Presence	Relative mobility	Molecular weight (KD)
1	29	0.27	127.35
2	14	0.30	118.35
3	8	0.31	112.71
4	29	0.34	104.76
5	4	0.36	99.77
6	29	0.39	92.72
7	28	0.41	86.17
8	6	0.44	80.09
9	29	0.52	64.29
10	29	0.56	58.31
11	29	0.60	52.89
12	29	0.62	49.15
13	14	0.66	44.58
14	23	0.68	42.46
15	27	0.70	40.43
16	12	0.71	38.51
17	29	0.74	35.79
18	29	0.77	33.29
19	29	0.80	30.17
20	28	0.86	26.06
21	29	0.90	23.63
22	29	0.94	20.92
23	29	0.98	18.97

Table 4 - Average cluster analysis groups separately assessed properties

Group	Number genotype	Weight of 100grains	Biomass	Secondary branches	Primary branches	Seeds in 100 pods	Number of empty pods	Number of full pods	Harvest index	Height lowest pod	Height	Days to reach	Yield	Hook sizes	Number of hooks	Flower development	Germination
1	7	6.4	832	7	3.9	239.3	12.8	44.2	51.7	12.6	40.1	198.1	218.1	1.9	2.1	76.3	84.6
2	10	6.7	753.4	6.5	3.7	239.4	8.2	38.6	82.2	11.8	38.1	192.6	192.8	1.8	2.4	78.6	79.8
3	12	6.4	872.4	8	4.2	259.1	7.5	48.9	54.1	13.8	42.6	213.7	219.8	2	2.4	76.8	81.9
Total	-	6.5	821.6	7.2	3.9	247.5	9.5	44.2	63.2	12.8	40.5	202.6	210.0	1.9	2.3	77.3	81.8

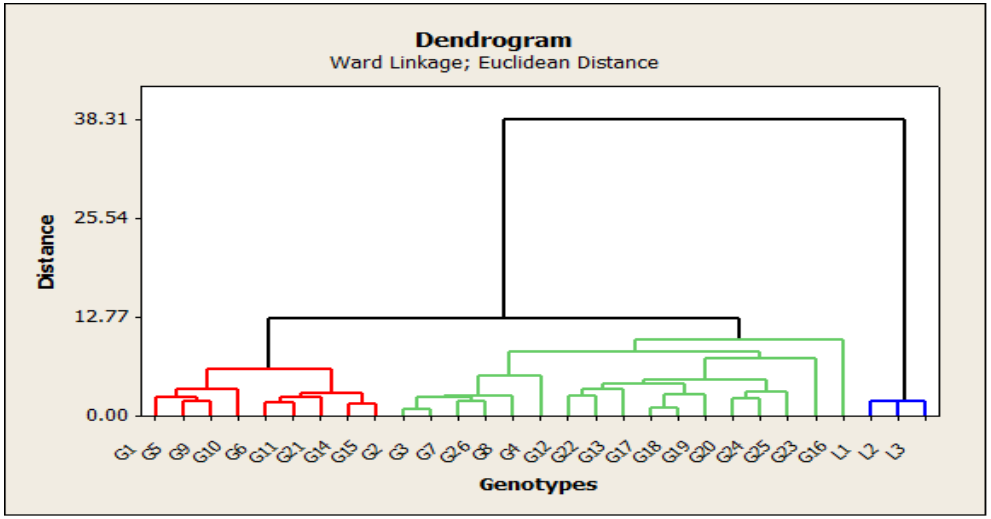

Figure 1: Dendrogram resulting from cluster analysis of minimum variance method (ward) in the genotypes studied and evaluated based on the properties

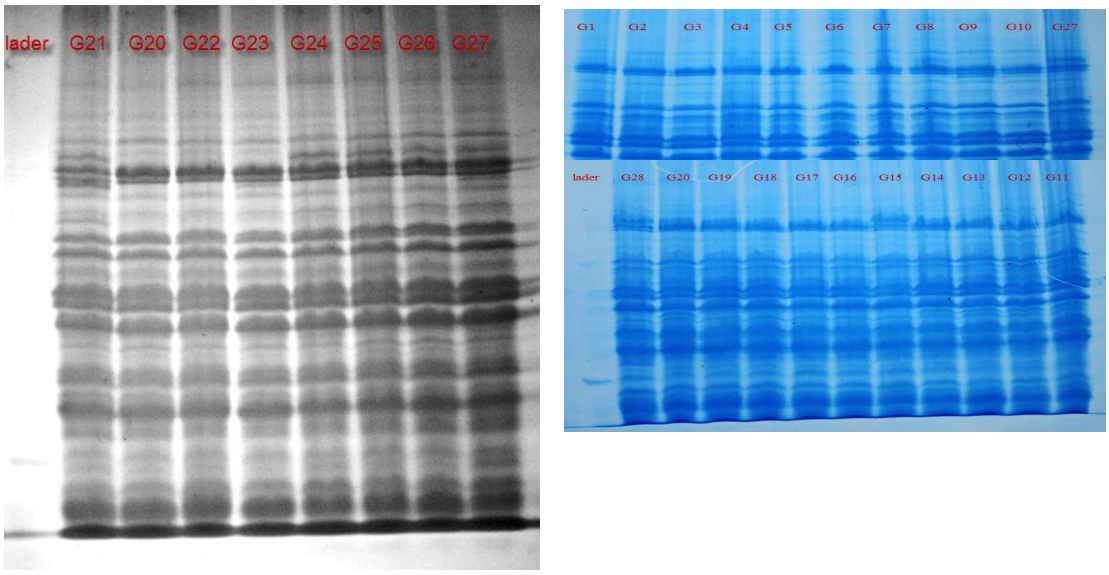

Figure 2: The pattern of protein bands

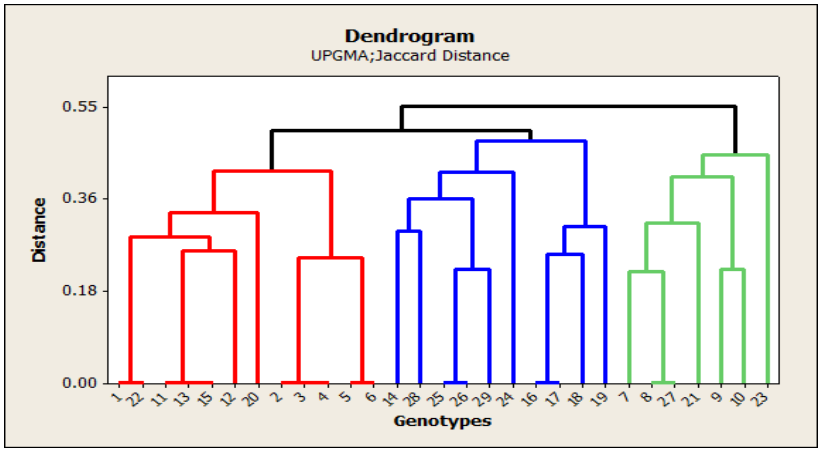

Figure 3: Dendrogram derived from UPGMA cluster analysis in lentil genotypes based on electrophoretic banding patterns

	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12	G13	G14	G15	G16	G17	G18	G19	G20	G21	G22	G23	G24	G25	G26	L1	L2	L3
G1	1																												
G2	0.947	1																											
G3	0.947	1	1																										
G4	0.947	1.000	1.000	1.000																									
G5	0.900	0.950	0.950	0.950	1.000																								
G6	0.900	0.950	0.950	0.950	1.000	1.000																							
G7	0.857	0.905	0.905	0.905	0.952	0.952	1.000																						
G8	0.818	0.864	0.864	0.864	0.909	0.909	0.955	1.000																					
G9	0.900	0.857	0.857	0.857	0.818	0.818	0.864	0.826	1.000																				
G10	0.947	0.900	0.900	0.900	0.857	0.857	0.905	0.864	0.950	1.000																			
G11	0.944	0.895	0.895	0.895	0.850	0.850	0.810	0.773	0.850	0.895	1.000																		
G12	0.895	0.850	0.850	0.850	0.900	0.900	0.857	0.818	0.810	0.850	0.944	1.000																	
G13	0.944	0.895	0.895	0.895	0.850	0.850	0.810	0.773	0.850	0.895	1.000	0.944	1.000																
G14	0.842	0.800	0.800	0.800	0.762	0.762	0.727	0.773	0.762	0.800	0.889	0.842	0.889	1.000															
G15	0.944	0.895	0.895	0.895	0.850	0.850	0.810	0.773	0.850	0.895	1.000	0.944	1.000	0.889	1.000														
G16	0.842	0.895	0.895	0.895	0.850	0.850	0.810	0.773	0.762	0.800	0.889	0.842	0.889	0.889	0.889	1.000													
G17	0.842	0.895	0.895	0.895	0.850	0.850	0.810	0.773	0.762	0.800	0.889	0.842	0.889	0.889	0.889	1.000	1.000												
G18	0.800	0.850	0.850	0.850	0.900	0.900	0.857	0.818	0.727	0.762	0.842	0.895	0.842	0.842	0.842	0.944	0.944	1.000											
G19	0.789	0.842	0.842	0.842	0.800	0.800	0.762	0.727	0.714	0.750	0.833	0.789	0.833	0.833	0.833	0.941	0.941	0.889	1.000										
G20	0.889	0.842	0.842	0.842	0.800	0.800	0.762	0.727	0.800	0.842	0.941	0.889	0.941	0.833	0.941	0.833	0.833	0.789	0.778	1.000									
G21	0.818	0.864	0.864	0.864	0.826	0.826	0.870	0.913	0.909	0.864	0.773	0.739	0.773	0.773	0.773	0.773	0.773	0.739	0.727	0.727	1.000								
G22	1.000	0.947	0.947	0.947	0.900	0.900	0.857	0.818	0.900	0.947	0.944	0.895	0.944	0.842	0.944	0.842	0.842	0.800	0.789	0.889	0.818	1.000							
G23	0.850	0.810	0.810	0.810	0.857	0.857	0.818	0.783	0.857	0.810	0.800	0.850	0.800	0.714	0.800	0.714	0.714	0.762	0.667	0.750	0.783	0.850	1.000						
G24	0.800	0.850	0.850	0.850	0.810	0.810	0.773	0.818	0.727	0.762	0.842	0.800	0.842	0.842	0.842	0.842	0.842	0.800	0.789	0.789	0.818	0.800	0.762	1.000					
G25	0.850	0.810	0.810	0.810	0.857	0.857	0.818	0.864	0.773	0.810	0.895	0.947	0.895	0.895	0.895	0.800	0.800	0.850	0.750	0.842	0.783	0.850	0.810	0.850	1.000				
G26	0.850	0.810	0.810	0.810	0.857	0.857	0.818	0.864	0.773	0.810	0.895	0.947	0.895	0.895	0.895	0.800	0.800	0.850	0.750	0.842	0.783	0.850	0.810	0.850	1.000	1.000			
L1	0.818	0.864	0.864	0.864	0.909	0.909	0.955	1.000	0.826	0.864	0.773	0.818	0.773	0.773	0.773	0.773	0.773	0.818	0.727	0.727	0.913	0.818	0.783	0.818	0.864	0.864	1.000		
L2	0.842	0.800	0.800	0.800	0.850	0.850	0.810	0.773	0.762	0.800	0.889	0.944	0.889	0.889	0.889	0.889	0.889	0.944	0.833	0.833	0.696	0.842	0.800	0.750	0.895	0.895	0.773	1.000	
L3	0.810	0.773	0.773	0.773	0.818	0.818	0.783	0.826	0.818	0.773	0.850	0.900	0.850	0.850	0.850	0.762	0.762	0.810	0.714	0.800	0.826	0.810	0.857	0.810	0.950	0.950	0.826	0.850	1.000

Table 5: Jaccard similarity coefficient based on the studied genotypes

Conclusion:

According to cluster analysis results, the first group with genotypes of 1, 5, 6, 9, 10, 11, 14, 15 and 21was the best group.

The most remote distance on protein bands was related to the genotype numbers of 12 and 14, 17, 18 and 19.

The highest number of bands (22) was observed in genotypes of 8 and 21 and control genotype of 27 and the least number of bands (16) was observed in genotypes of 19, and 20.

Bands number of 2, 3, 5, 7, 8, 13, 14, 15 and 16 showed polymorphisms with molecular weight of 118.38, 112.71, 77.99, .1786, 80.09, 58.48, 46.42, 43.40, and 51. 38, respectively.

The third group included 12 genotypes of G1, G22, G11, G13, G15, G12, G20, G2, G3, G4, G5 and G6 with highest number of members which were late crop, high yielding and legged. Also, they obtained high values in studied traits.

categorizations based on morphological traits and protein bands had a consistency at 35%

References

1- Bozorgi, A. 1995. DNA molecular markers in plant breeding. Journal of Construction Research, No. 22, Spring 1373.

2- Huff, D. R., R. Peakall, and Smouse. P. E. 1993. RAPD variation within and among natural populations of outcrossing buffalograss [Buchloë dactyloides (Nutt.) Engelm.]. Theor. Appl. Genet. 86: 927–934. Huff, D.R., 1997. RAPD characterization of heterogeneous perennial ryegrass cultivars. Crop Science, 37: 557-594.

3- Jaccard, P. 1908. Nouvelles researches sur la distribution florale. Bull. Soc. Vaudoise Sci. Natl. 44: 223- 270.

4- Ladizinsky, G., and Hymowitz, T. 1979.Seed Protein Electrophoresis in Taxonomicand Evolutionary Studies. Theor. Appl.Genet., 51:145-151.

5- Liu, Z-H., Jarret, R L., Duncan, R R and Kresovich, S. 1994. Genetic relationship and variation among ecotypes of seashore paspaium (Paspalum vaginaturn) determined by random amplifiecl polymor pHic DNA markers. Genome 37: 101 1-1017.

6- Moghaddam, M, Mohammadi, SA. and Aghaei, M. 1995. Introduction to multivariate statistical methods (translation). Leading publisher of science.

7- Mousavi Zadeh, SA. 2007. Study of genetic diversity of morphological and molecular markers using onions Iran. PhD thesis, Faculty of Agriculture, Tabriz University.

8-Peakall, R, Smouse, P. Eo and Huff, D. R. 1995. Evolutionary implications of allozyme and RADP variation in diploid populations of dioecious buffalograss Buchloe dactyloides. Molecular Ecology, 4: 135-147.

9-Rahimi Nejad, M., 1995, Plant Molecular systematic, Isfahan University Press, 506 pages.

10-Sahai, S. and Rana, R.S. 1977. Seed proteinhomology and elucidation of species relationships in Phaseolus and Vigna species. New Phytol., 79: 527-34.

11-Shehata, M.M. 2004. The roles of seed proteins and RAPD-PCR in genotyping variabilities of some wheat (Triticum vulgare L.) cultivars. Pak. J. Biol. Sci., 7: 984-994.

12-Wu, L. and Lin.IL.1994. Identimg buffalograss (Buchloe dactyloides Nutt. Engelrn) Cultivar breeding b e s using random amplified polymor pHic DNA (RAPD) markers. J. Am. Soc. Hort. Sci. 119: 126-130.