INNSPUB

International Journal of Biosciences | IJB |
ISSN: 2220-6655 (Print) 2222-5234 (Online)
http://www.innspub.net
Vol. 3, No. 5, p. 101-107, 2013

RESEARCH PAPER OPEN ACCESS

Evaluation of seed storage proteins in common bean by some biplot analysis

Akbar Marzooghian[1*], Mostafa Valizadeh[1], Mohammad Moghaddam[1], Mohammad Hasan Kooshki[2]

[1]*Department of Plant Breeding and Biotechnology, University of Tabriz, Tabriz, Iran*

[2]*Agricultural and Natural Resources Research Center of Lorestan Province, Broujerd, Iran*

Key words: Cluster analysis, GGE biplot, factor analysis, *Phaseolus vulgaris* L., seed storage proteins.

doi: http://dx.doi.org/10.12692/ijb/3.5.101-107 Article published on May 20, 2013

Abstract

In order to study of seed storage proteins, proteins samples of common bean genotypes were prepared by 0.2 M NaCl of extracting soluble. Genotypes were located in two groups by cluster analysis using Wilks' lambda statistic. Two groups were different for yield components (number of pods per plant, number of seeds per plant and seed weight). Factor analysis showed that two factors described 61% of total proteins variation. Correlated bands with yield components characters had the highest coefficients for the first factor. This factor was named "yield components proteins". Protein bands via RM 58 and 64 had relationship with days to flowering. Therefore, the second factor was named "phenologic proteins". Genotypes were located in four groups by these factors. Length, angle and presence of protein bands were important characteristics to explain graphical information in GGE biplot compared to factor analysis.

* **Corresponding Author:** Akbar Marzooghian ✉ marzooghian@tabrizu.ac.ir

Introduction

The results of SDS-PAGE have been used correctly to evaluation of between genus genetic diversity (Ocompo and Toro, 2008; 2004; Garvin and Weeden, 1994) and species (Duran *et al.*, 2005). The banding patterns produced by seed protein electrophoresis have been used to effectively characterize cultivars of pasture grasses and legumes (Sheidai *et al.*, 2000). Multiple domestication centers have been suggested through the seed storage protein electrophoresis analysis from different wild and cultivated accessions of common bean (Gepts *et al.*, 1986). Seed storage protein electrophoresis has been also used to estimate diversity among accessions in genetic resources collection (Gardiner *et al.*, 1992). Generally, proteins bands such as seed storage proteins have been used as markers in the following four main applications: analysis of genetic diversity within and among populations (Marzooghian *et al.*, 2011; Gepts, 1990b), plant domestication in relation to genetic resources conservation and breeding, genome relationships, and as a tool in plant breeding (Gepts, 1990b). Diversity in the types of phaseolin, the major seed storage protein in common bean, has been especially useful for classifying beans into Andean and Mesoamerican gene pools since most of the cultivars from one center of domestication possess a certain set of phaseolin types which are not found in cultivars or wild types from the other center of domestication (Gepts, 1990a).

Common bean (*Phaseolus vulgaris* L.) is the most important edible food legume in the world, representing 50% of the grain legumes for direct human consumption (McClean *et al.*, 2004). China, Iran, Turkey, and Japan are the most important countries that produce common bean in Asia. Common bean has the highest yield than other food legumes in Iran (FAO, 2003). Three types of white, red and pinto bean are produced in Iran.

Some appropriate methods such as cluster analysis, PCA, factor analysis and GGE biplot are used for genetic diversity evaluation, parental selection, study interaction between the genotypes and environments and applications to other types of two-way data (Aharizad *et al.*, 2012; Eivazi *et al.*, 2008; Mohammadi and prasanna, 2003; Bhatt, 1970) When a large number of variables had relationship, factor analysis transforms these variables to smaller number of unobservable factors. A method widely used for determining a first set of loadings is the principal component method. This method seeks values of the loadings that bring the estimate of the total communality as close as possible to the total of the observed variation (Walton, 1971). Yan (2001) provided to the agricultural research community an excellent scientific method of visual analysis, called GGE biplot analysis.

Plants' choice is the first step in plant breeding program to hybridization. In order to benefit transgressive segregation, genetic distance between parents is necessary (Joshi *et al.*, 2004). Also, Recombination and selection methods depend mainly upon the genetic distance among parents, breeding objectives and available resources. Maintenance and availability of germplasm as a source of genetic variation is especially important to fulfill the increasing needs of breeders. The objective of the study was to use seed storage proteins to study variety inter-relationships and the role of these proteins to selection among genotypes of *P. vulgaris* by formal statistical and graphical analysis.

Materials and methods

Seventy common bean genotypes randomly selected from collection genotypes exist in Iran (data not shown) were evaluated it this study. The common bean genotypes were obtained from National Bean Research Station of Khomeyn, Iran.

Protein patterns were studied by SDS-PAGE. The method of Krochko and Bewley was used for the extraction of soluble seed storage proteins in salt (Krochko and Bewley, 2000). Low salt (0.2 M NaCl) solution was used in this research. After seed coat separation, seeds were ground and the resulting flour was filtered by a sieve (40 mesh). Forty mg of floured

seed was poured in a micro tube. Then extraction solution was added in each micro tube and soluble protein samples in low salt were prepared. Polyacrylamide gels and buffers were prepared by Hames and Richwood method (Hames and Richwood, 1990). The Laemmli method was used for protein electrophoresis (Laemmli, 1970). Electrophoresis was performed using vertical gels (10%) with 20 µl loading (Table 1). After staining, protein bands were evaluated qualitatively. Each band was named according to its relative mobility (RM). A zero-one coding was used for the presence or absence of proteins in a special location.

Furthermore, the genotypes under study were evaluated for several agronomic characters such as number of days to flowering, number of days to maturity, plant height, pod number per plant, seed number per plant and 100 seed weight in National Bean Research Station of Khomeyn, Iran. Seed length, width and thickness were also measured for three grains of each genotype. Relationship between agronomic characters and protein bans was calculated by t-test statistic. UPGMA base on simple maching coefficient was used for genotypes clustering. Discriminate analysis based on Wilks' lambda (Wilks' lambda = $SS_{within\ groups}/SS_{total}$) was used to identify cutting point in cluster tree. Factor and GGE biplot analysis was also carried out to explain the variation.

Statistical analyses were performed by SPSS, STATISTITA and GGE biplot software.

Results

Cluster analysis

Electropherogram of several common bean genotypes in terms of soluble proteins based on the method of extraction are shown (Fig. 1). Cluster tree was cut via discriminate analysis using Wilks' lambda statistics (Table 2) and consequently, genotypes were located in two groups (Fig. 2). Genotypes that located in group1 had higher number pod per plant, number seed per plant and seed weight than genotypes in other group (Table 2).

Table 1. Consumed materials for the preparation of storage protein sample and loading of sample in common bean.

Sample type	Loaded sample in wells (µl)	2–mercaptoethanol (µl)	Reload buffer (µl)	Extracted soluble protein (µl)
S_1	20	2.5	7	15

S_1 = Soluble proteins in low salt.

Table 2. Group means difference of studied common bean genotypes by T-test in cluster analysis.

		Characteristics		
	Groups	Pod number per plant	Seed number per plant	Seed weight
Mean	1	12.95	45.37	44.17**
	2	19.89**	76.42**	32.38
Wilks' Lambda	0.036**			

** Significant at the 1% level of probability

Biplot analysis

Factor analysis transformed electrophoresis bands into two factors (Table 3). These factors explained 61% of total for proteins variation. First factor described 48% of the variation. Proteins via RM 17, 18, 30, 32, 38, 40 and 54 had the highest coefficients for first factor. Genotypes located in two groups for this factor (Fig.3). Comparing two groups it was revealed that one of groups had lower 100 seed weight and higher pod number per plant and seed number per plant than the other group (Table 4). Therefore, this factor was named "yield components

proteins". Furthermore, the genotypes were separated into two groups for the second factor. These groups were also different for days to flowering. Thus, this factor was named "phenologic proteins". Considering the independence of factors for "yield components proteins" and "phenologic proteins", it seems these proteins bands could be used to select genotypes simultaneously for the above agronomic and phenologic characters.

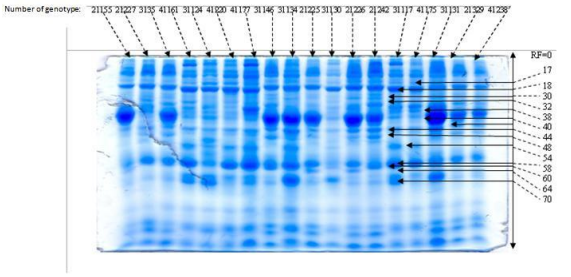

Fig. 1. Gel samples of several common bean genotypes for the extraction method of soluble proteins in low salt.

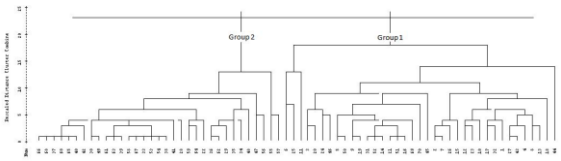

Fig. 2. UPGMA dendrogram based on simple matching coefficient showing relationship among 70 studied common bean based on electrophoresis bands.

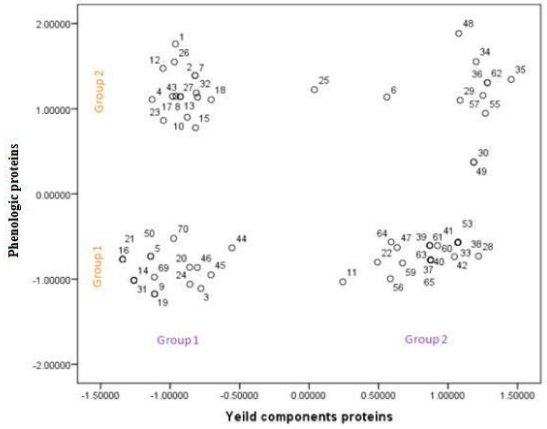

Fig. 3. Features of studied common bean genotypes based on their factor scores.

Correlations between genotypes and protein bands have been shown in figure 4 by GGE biplot analysis. Genotypes classification in biplot figures (Figs. 3 and 4) was similar, because of principle components method (PC) was used in both GGE biplot and factor analysis.

Table 3. Factor analysis based on principal component analysis of protein bands in studied common bean genotypes.

	Components		
Proteins via RM	Factor 1	Factor 2	Communality
3	0.644	-0.154	0.486
5	0.312	0.075	0.836
7	-0.632	-0.243	0.461
11	0.385	-0.198	0.710
13	-0.171	0.307	0.678
17	-0.920	-0.092	0.856
18	0.920	0.092	0.856
30	0.945	0.150	0.943
32	0.945	0.150	0.943
38	-0.945	-0.150	0.943
40	-.0945	-0.150	0.943
48	-0.748	0.023	0.591
54	0.869	-0.022	0.765
58	-0.250	0.882	0.940
60	0.301	-0.545	0.516
64	-0.322	0.866	0.938
70	0.614	0.023	0.436
Variance	48.835%	12.681%	
Cumulative variance (%)	48.835%	61.516%	

Factor 1 = Yield components proteins; Factor 2 = Phenologic proteins

Discussion

The amount of graphical information in GGE biplot was more than the factor analysis because of presence, length and angle of vectors of proteins bands. Effective protein bands can be identified with vector length. Proteins bands 17, 18, 30, 32, 38, 40, 54 for principle component 1(PC1) and proteins bands 58 and 64 for PC2 were located between two groups for each principle component. A protein

bands located near the biplot origin has little effect on genotypes grouping.

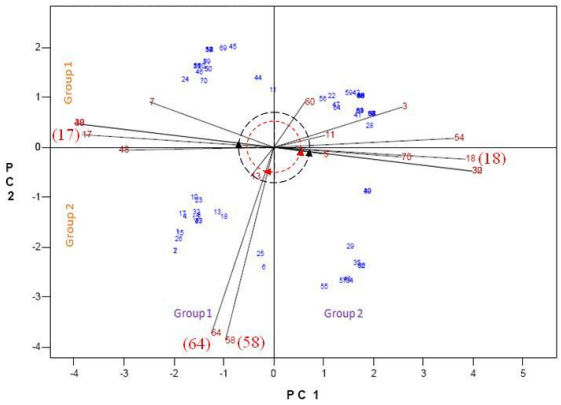

Fig. 4. Features of studied common bean genotypes based on their PC scores in GGE biplot analysis.

Amount of bands correlation can be identified by their angle. If the angle is 90˚, the bands are independent and can be concluded loci for these bands are probably different. It seems that bands via RM 58 and 64 had no relationship with mentioned proteins bands in PC1. In the other hand, bands had 180˚ angle they had different control for correlated trait. Protein bands with this angle more probably are alleles with each other, for example Protein bands 17 and 18.

Table 4. T-test analysis for groups in yield components proteins and phenologic proteins factors.

	yield components proteins			Phenologic proteins		
Characteristics	Groups	Mean	Sig.	Groups	Mean	Sig.
Seed thickness	1	0.6104	0.166	1	0.5907	0.617
	2	0.2822		2	0.6010	
Seed length	1	1.2762	0.398	1	1.2715	0.173
	2	1.3097		2	1.3256	
Seed width	1	0.8052	0.997	1	0.8045	0.943
	2	0.8051		2	0.8061	
Days to flowering	1	47.28	0.405	1	45.22	0.008
	2	46.17		2	48.69	
Days to maturity	1	98.00	0.635	1	96.44	0.551
	2	96.55		2	98.26	
Plant height	1	63.03	0.103	1	54.29	0.193
	2	52.50		2	62.90	
Pod number per plant	1	20.17	0.010	1	17.17	0.335
	2	12.95		2	14.38	
Seed number per plant	1	78.76	0.004	1	65.04	0.265
	2	45.37		2	51.17	
Seed weight	1	32.28	0.000	1	39.41	0.690
	2	44.47		2	38.34	

One application of evaluation for diversity is to choose genotypes from two ends of the phenotypic distribute on. Graphical information obtained from biplot analysis was more than the Cluster analysis. For instance, bands 58 and 64 had no effect to group separation, while these bands had main role to genotypes classification in biplot analysis. Crossing of the genotypes in the opposite locations in the distribution allows the breeders to increase the probability of heterosis and transgressive

segregation. Significant heterosis has also been found for number of days to flowering (Barelli *et al.*, 2000; Mitranov, 1983), plant height (Gonçalves-Vidigal *et al.*, 2008), number of pods per plant, number of seeds per plant, seed weight (Gonçalves-Vidigal *et al.*, 2008; Barelli *et al.*, 2000; Nienhuis and Singh, 1988) seed thickness, seed length and seed width (Corte *et al.*, 2010) in beans. The results of analysis pointed above can be used to breeding programs.

Conclusion

Common bean genotypes were located in different groups based on seed storage proteins. Selection of genotypes in these groups can help breeders to indirect selection for some traits, accumulate favorable alleles and broaden the genetic base. Genotypes can be selected based on factor scores or PC scores and improvement for several traits, simultaneously. GGE biplot was better graphical tool than factor analysis because of present, angel and length of protein bands. Graphical analysis and formal statistical analysis are complementary to maximize understanding of the data.

Acknowledgements

Authors thank to Hamid Reza Dorri for assistance in evaluating genotypes in field experiments. This research was supported by University of Tabriz.

References

Aharizad S, Sabzi M, Mohammad SA, Khodadadi E. 2012. Multivariate analysis of genetic diversity in wheat (*Triticum aestivum* L.) recombinant inbred lines using agronomic traits. Annals of Biological Research **3**, 2118-2126.

Barelli MAA, Celeste M, Vidigal Filho PS, Scapim CA. 2000. Combining ability among six common bean cultivars adapted to the North West Region of Paraná State, Brazil. Bragantia **59**, 159-164 http://dx.doi.org/10.1590/S0006-87052000000200006

Bhatt GM. 1970. Multivariate analysis approach to selection of parents for hybridization aiming at yield component in self pollination crops. Australian Journal of Agricultural Research **21**, 1-7.

Corte AD, Moda-Cirino V, Arias CAA, Toledo JFF, Destro D. 2010. Genetic analysis of seed morphological traits and its correlations with Grain yield in common bean. Brazillian Archives of Biology and Technology **53**, 27-34. http://dx.doi.org/10.1590/S1516-89132010000100004

Durán L, Blair M, Giraldo M, Machiavelli R, Prophete E, Nin J, Beaver J. 2005. Morphological and molecular characterization of common bean (*Phaseolus vulgaris* L.) landraces from the Caribbean. Crop Science **45**, 1320-1328.

Eivazi R, Naghavi MR, Hajheidari M, Pirseyedi SM, Ghaffari MR, Mohammadi SA, Majidi I, Salekdeh GH, Mardi M. 2008. Assessing wheat (*Triticum aestivum* L.) genetic diversity using quality traits, amplified fragment length polymorphisms, simple sequence repeats and proteome analysis. Annals of Applied Biology **152**, 81–91. http://dx.doi.org/10.1111/j.1744-7348.2007.00201.x

FAO. 2003. Production Yearbook. Vol. 56. *Rome*.

Gardiner SE, Forbe MB. 1992. Identification of cultivars of grasses and forage legumes by SDS-PAGE of seed protein. In Seed analysis. (H. F. Linskens and J. F. Jackson, ed.), Springer-Verlag, Berlin, New York, 43–61

Garvin DF, Weeden NF. 1994. Isozyme evidence supporting a single geographic origin for domesticated tepary bean. Crop science **34**, 1390-1395.

Gepts P. 1990 a. Biochemical evidence bearing on the domestication of *Phaseolus* (Fabaceae) beans. Economic Botany **44**, 28-38.

Gepts P. 1990 b. Genetic diversity of seed storage proteins in plants. In Plant populations genetics, breeding and genetic resources. (A. H. D. Brown, M.T. Clegg, A. L. Khaler and B. S. Weir, ed.), Sinauer Associates Inc., Sunderland, Massachusetts, 64–82.

Gepts P, Osborn TC, Rashka K, Bliss FA. 1986. Phaseolin-protein variability in wild form and landraces of the common bean (*Phaseolus vulgaris*); evidence for multiple centres of domestication. Economic Botany **40**, 451–468.

Gonçalves-Vidigal MC, Silvero L, Elias HT, Filho PSV. 2008. Combining ability and heterosis in common bean cultivars. Pesquisa agropecuaria brasileira **43**, 1143-1150.

Hames B, Richwood D. 1990. Gel electrophoresis of proteins: a practical approach. Oxford University Press, USA.

Joshi BK, Mudwari A, Bhatta MR, Ferrara GO. 2004. Genetic diversity in Nepalese wheat cultivars based on agromorphological traits and coefficients of parentage. Nepal Agriculture Research Journal **5**, 7-17.

Krochko JE, Bewley DJ. 2000. Seed storage proteins in cultivars and subspecies of alfalfa (*Medicago sativa* L.). Seed Science Research **10**, 423-434.
http://dx.doi.org/10.1017/S0960258500000477

Laemmli UK. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. nature **227**, 680-685.

Marzooghian A, Valizadeh M, Moghaddam M, Kooshki MH. 2011. The comparison of genetic diversity between two extraction methods of seed storage proteins in common bean. Clinical Biochemistry **44**, 256-256.
http://dx.doi.org/10.1016/j.clinbiochem.2011.08.640

McClean P, Kami J, Gepts P. 2004. Genomics and genetic diversity in common bean. Legume Crop Genomics. AOCS Press, Champaign, 60–82.

Mitranov L. 1983. A study of general and specific combining ability for productivity in kidney bean (*Phaseolus vulgaris* L.) cultivars. Genetics and Plant Breeding (Sofia) **16**, 176-180.

Mohammadi SA, Prasanna BM. 2003. Analysis of Genetic Diversity in Crop Plants—Salient Statistical Tools and considerations. Crop Science **43**, 1235-1248.

Nienhuis J, Singh S. 1988. Genetics of Seed Yield and its Components in Common Bean (*Phaseolus vulgaris* L.) of Middle-American Origin. Plant breeding **101**, 155-163.

http://dx.doi.org/10.1111/j.1439-0523.1988.tb00281.x

Ocompo CH, Toro CO. 2004. Biochemical evidence supporting the existence of a weedy form in tepary bean (*Phaseolus acutifolius* A. Gray). Bean Improvement Cooperative. Annual report **47**, 165-166.

Ocompo CH, Toro CO. 2008. Phaseolin diversity of Nicaraguan common bean germplasm held at CIAT). Bean Improvement Cooperative. Annual report **51**, 122-123.

Sheidai M, Hamta A, Jaffari A, Noori-Daloli MR. 2000. Morphometric and seed protein studies of *Trifolium* species and cultivars in Iran. Plant Genetic Resources Newsletter **120**, 52–54.

Walton PD. 1971. The use of factor analysis in determining characters for yield selection in wheat. Euphytica **21**, 416-421.
http://dx.doi.org/10.1007/BF00035667.

Yan W. 2001. GGEBiplot—A Windows application for graphical analysis of multi-environment trial data and other types of two-way data. Agronomy Journal **93**, 1111–1118.
http://dx.doi.org/10.2134/agronj2001.9351111x