IJST (2015) 39A4: 567-572
Iranian Journal of Science & Technology

http://ijsts.shirazu.ac.ir

Genotype profiling of seed storage proteins in wheat (*Triticum aestivum* L.)

N. Tahir[1], S. Ali[1], A. Ijaz[1], Mahmood-ur-Rahman Ansari[1]* and N. Iqbal[2]

[1]*Department of Bioinformatics and Biotechnology, GC University, Allama Iqbal Road, Faisalabad-38000, Pakistan*
[2]*Nuclear Institute for Agriculture and Biology (NIAB), Jhang Road, Faisalabad, Pakistan*
E-mail: mahmoofansari@gcuf.edu.pk

Abstract

Seed–storage proteins are important reservoirs of food and energy which are also involved in the determination of bread making quality. Solubility properties of these proteins are traditionally classified into four classes: albumin, globulin, prolamin and glutelin. Gliadin and glutenin have also been studied extensively and the genetics and biochemistry are relatively well known. Grain proteins from 18 Pakistani wheat genotypes were checked for genetic diversity evaluation based on 15% sodium dodecylsulphate polyacrylamide gel electrophoresis (SDS–PAGE). According to quality score, WL–711, Abadghar, Yacoora, Punjab–96 and Zardana showed the highest score of 10. All these varieties had subunit 5+10 which had positive correlation for bread-making quality. Different Bioinformatics tools like STRUCTURE, TASSEL AND STATISTIXL were used to examine the data obtained from different varieties and the outcomes were conveyed in the course of a dendrogram revealing the conflicts and similarities among varieties. STRUCTURE was used for the analysis of population structure. This tool performed the grouping of 18 wheat varieties into two groups. TASSEL was used for cluster analysis, which grouped the varieties under study into two main groups containing the same individuals as described by other tools. STATISTIXL performed a cluster analysis using the similarity measure of Jaccard. The results were shown in the form of dendrogram on the basis of group average. Dendrogram revealed two main clusters differentiated as A and B at the distance of about 85%. With the aim of improving flour quality, wheat seed storage protein fingerprinting is used to govern the gluten protein pattern in studies.

Keywords: Storage Proteins; wheat; glutelin; SDS–PAGE; dendrogram; Bioinformatics

1. Introduction

Wheat (*Triticum* spp.) is a grass, cultivated worldwide for food. In Pakistan, during the fiscal year 2011–12, wheat contributed 12.5% of value added in agriculture and 2.6% to GDP (Economic Survey of Pakistan, 2011–12). Wheat grain is a staple food used to make flour for breads, biscuits, cookies, cakes, breakfast cereals, pasta, and noodles. Wheat is planted to a limited extent as a forage crop for livestock, and its straw can be used as a construction material for roofing slate (Farooq et al., 2010).

Storage proteins of wheat seeds (*Triticum aestivum* L.) are important reservoirs of energy and food, and also deal with the determination of the bakery quality (Shewry and Halford, 2002). The quality parameters are strongly correlated with the SDS or acetic acid extractable glutenin. Quality relationships are weaker for high molecular glutenin subunits. In some studies, it was shown that the presence of some high glutenin subunit molecular correlate with the amount of non-extractable glutenin. Therefore, the subunits are probably ultimately related to the quality of bread by the amount of non-extractable glutenin (Zilic et al., 2011).

Small scale testing has recently been set up to analyze small amounts of glutenin fractions. The importance of high molecular glutenin has been shown by controlling breeding studies and, to a lesser extent, for the quality. Many of the resolutions have not described the amount of glutenin (Huang and Khan, 1997). Seed protein electrophoresis is increasingly being used as an additional method for the documentation of species, as well as a useful tool for the follow-up of the evolution of different groups of plants. Using the protein profile of the seed obtained by the resolution of taxonomic and evolutionary problems has expanded considerably over the last decade. It has been applied in more than 45 different genera belonging to 13 families of plants. A number of researchers have studied the importance of the seed protein profile taxonomic and evolutionary goals (Singh et al., 2007).

*Corresponding author
Received: 23 August 2013 / Accepted: 8 October 2014

Constancy is one of the central features of seed protein profiles. The regularity of the character profile additive and seed protein electrophoresis is unique and they are powerful tools to examine the source and evolution of polyploid plants. The seed protein composition is highly stable. Mature seeds of different ages that have the same profile as the proteins are mainly seed storage proteins and are not likely to be changed in the mature dry seed (Tahir et al., 1995). Morphological markers were not enough to describe the morphological genetic diversity among cultivars and accessions. So, this study was designed to evaluate the genetic diversity of seed storage proteins of local wheat genotypes. The genetic variability observed could be useful in any selection strategy (Omid and Valizadeh, 2009).

2. Materials and Methods

2.1. Plant Samples

Eighteen wheat genotypes were obtained from a collection of wheat germplasm currently available with Marker assisted Breeding Group, Plant Breeding and Genetics Division, Nuclear Institute for Agriculture and Biology, Faisalabad, Pakistan. The samples were comprised of varieties grown in different ecological regions of Pakistan, as well as some exotic lines. The wheat varieties included in the present work are presented in Table 1. Three reference varieties (with known subunits of HMW-Glu) were used to detect different subunits in the genotypes under study and are presented in Table 2.

2.2. Protein Extraction

Random samples of three seeds from each genotype were selected for total protein extraction. The seeds were ground to fine powder with pestle and mortar. Ethanol solution (70%) was added.

Table 1. Local wheat varieties used in study

S. No.	Wheat Varieties	Reference No.	Origins
1	HD-2329	7	Punjab
2	C-217	8	Sindh
3	C-228	9	Punjab
4	WL-711	11	Punjab
5	M-H-9-7	12	Punjab
6	Inqlab-91	16	Punjab
7	Marwat	17	Punjab
8	Sulman-96	24	NWFP
9	Abadghar	31	Sindh
10	Kiran	32	Sindh
11	Iqbal-2000	36	Punjab
12	Saleem-2000	41	NWFP
13	AS-2002	46	Punjab
14	SH-2002	47	Punjab
15	Yacoora	48	Sindh
16	Punjab-96	51	Punjab
17	Paban	54	Sindh
18	Zardana	62	Balochistan

Table 2. Reference varieties used for protein profiling and their subunits

S. No.	Varieties	HMW Subunits
1.	Marquis	1, 7+9, 5+10
2.	Chinese Spring	Null, 7+8, 2+12
3.	Gabo	2*, 17+18, 2+12,

(400µl) to sample and kept for an hour at room temperature. The samples were vortexed after every 15 min. Then, samples were centrifuged for 10 minutes and the supernatant was discarded. At the end, 500µl sample buffer [0.5 M Tris-HCl (pH 6.8), 30% SDS, 10 % glycerol and 5% 2-mercaptoethanol] was added to each sample and held overnight at 40°C.

2.3. Protein Profiling

Seed composition was analyzed by SDS – PAGE slab type gel using 15% polyacrylamide. SDS – PAGE of total seed proteins was performed in a discontinuous buffer system according to the method of Laemmli, (1970). The gel was stained with Coomassie Brilliant Blue (CBB) and destained until the background became clear.

2.4. Data Analysis

For each variety, electropherogram was scored and the presence (1) or absence (0) of each subunit was observed. Presence and absence of data were entered into a binary data matrix (Rohlf, 2000).

3. Results

3.1. Genetic diversity evaluation

Eighteen wheat varieties used in this study showed different patterns of bands on SDS–PAGE. SDS–PAGE was performed in order to study the genetic diversity among wheat varieties. Electropherogram showed the protein banding pattern of different varieties of wheat (Fig. 1 & 2).

A total of 18 varieties were included in the study. Strains 1, 5 and 6, 4,9,17 and 10, 14 and 15, 16 and 7, 12 were common, while others showed significant variation (Table 3).

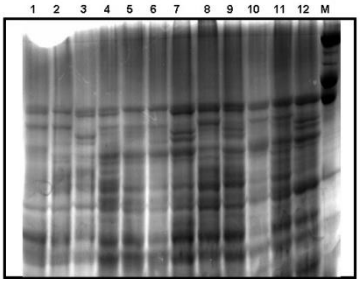

Fig. 1. Electrophorogram showing banding pattern of wheat proteins. L1-L12=Wheat Protein samples, and L-13=Protein marker

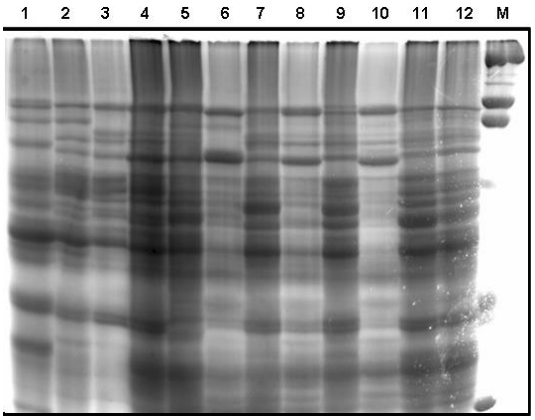

Fig. 2. Electrophorogram showing banding pattern of wheat proteins. L1-L12: Protein samples, L-13=Protein marker

3.2. Cluster analysis on the basis of SDS–PAGE

All of the commonly found allelic variants at the Glu-A1 locus (null, 1, 2) as described by Payne and Lawrence, (1983) were found in the genotypes under study. Frequency and partial frequency of 2* allele, i.e. 10 and 55.55 respectively, were high among all the alleles of a Glu-A1 locus. At Glu-B1 locus, 8 different HMW glutenin subunits (17+18, 7+9, 13+19, 15 and 20) appeared in Pakistani genotypes. Among these subunits 17+18 and 7+9 were the most frequent combinations. The frequency and partial frequency of subunit 17+18 are 11 and 22.22 respectively, while the subunit 7+9 had a frequency and partial frequency values of 4 and 61.11, respectively. The subunits 13+19 and 15 occurred least in this study with a frequency of 1 and partial frequency of 5.55 for each respectively.

At Glu-D1 locus only two subunit combinations were expressed, 5+10 and 2+12. The 2+12 subunit was the most dominant combination having the highest frequency and partial frequency values i.e.10 and 55.55 respectively (Table 4). A subunit compositions of Null, 17+18, 5+10 and 2*, 17+18, 2+12 were more frequent and occurred in 16.67% of the varieties. While 2*, 7+9, 2+12, 1, 17+18, 5+10 and 2*, 7+9, 5+10 subunit combinations occurred in 11.11% of the varieties. The third place, 6 different subunit combinations 2*, 17+18, 5+10, 1, 17+18, 2+12, 2*, 20/15, 2+12, Null, 20, 2+12, 2*, 13+19, 2+12, Null and 17+18, 2+12 exhibited the lowest frequency of 5.55% and each combination appeared in only one cultivar. The overall quality score for individual varieties ranged from 4 to 10 (Table 5). High scores of these varieties were due to the presence of subunits 2*,1, 17+18, 13+19 and 5+10, having the quality points of 4, 3, 3, 3, and 3 respectively.

The cluster analysis of 18 wheat varieties was performed using STATISTICA (Fig. 3). This tool performed a cluster analysis using the similarity measure of Jaccard. The results were shown in the form of dendrogram on the basis of group average.

Table 3. Distribution of seed storage protein subunits in different local wheat varieties

		Subunits																
		Glu-A1			**Glu-B1**										**Glu-D1**			
S. No.	**varieties**	Null	1	2*	6	7	8	9	13	14	15	17	18	19	20	2	5	10	12
1	HD-2329	0	0	1	0	1	0	1	0	0	0	0	0	0	0	1	0	0	1
2	C-217	0	0	1	0	0	0	0	0	0	1	0	0	0	1	1	0	0	1
3	C-228	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1
4	WL-711	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	0
5	M-H-9-7	0	0	1	0	1	0	1	0	0	0	0	0	0	0	1	0	0	1
6	Inqlab-91	0	0	1	0	0	0	0	0	0	0	1	1	0	0	1	0	0	1
7	Marwat	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0	1	1	0
8	Sulman-96	0	0	1	0	0	0	0	1	0	0	0	0	1	0	1	0	0	1
9	Abadghar	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	1	0
10	Kiran	0	0	1	0	0	0	0	0	0	0	1	1	0	0	1	0	0	1
11	Iqbal-2000	0	1	0	0	0	0	0	0	0	0	1	1	0	0	1	0	0	1
12	Saleem-2000	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0	1	1	0
13	AS-2002	1	0	0	0	0	0	0	0	0	0	1	1	0	0	1	0	0	1
14	SH-2002	0	0	1	0	0	0	0	0	0	0	1	1	0	0	1	0	0	1
15	Yacoora	0	1	0	0	0	0	0	0	0	0	1	1	0	0	0	1	1	0
16	Punjab-96	0	1	0	0	0	0	0	0	0	0	1	1	0	0	0	1	1	0
17	Paban	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	1	0
18	Zardana	0	0	1	0	0	0	0	0	0	0	1	1	0	0	0	1	1	0

Table 4. Frequency distribution of seed storage protein subunits

Glu-1 loci	**Glu-A1**			**Glu-B1**					**Glu-D1**	
Glu-1 alleles	Null	1	2*	7+9	17+18	13+19	15	20	2+12	5+10
Frequency	5	3	10	4	11	1	1	2	10	8
Partial frequency	27.78	16.67	55.55	22.22	61.11	5.55	5.55	11.11	55.55	44.44

Table 5. Combination of Glu-1 alleles and quality score of the varieties

S. No.	**Combination of Glu-1 Alleles**	**Percentage Frequency**	**No. of Cultivars**	**Quality Score**
1.	Null, 17+18, 5+10	16.67	3	8
2.	2*, 7+9, 2+12	11.11	2	7
3.	2*, 17+18, 2+12	16.67	3	8
4.	1, 17+18, 5+10	11.11	2	10
5.	2*,17+18, 5+10	5.55	1	10
6.	1,17+18, 2+12	5.55	1	8
7.	2*, 20/15, 2+12	5.55	1	6
8.	Null, 20, 2+12	5.55	1	4
9.	2*, 13+19, 2+12	5.55	1	8
10.	Null, 17+18, 2+12	5.55	1	6
11.	2*, 7+9, 5+10	11.11	2	9

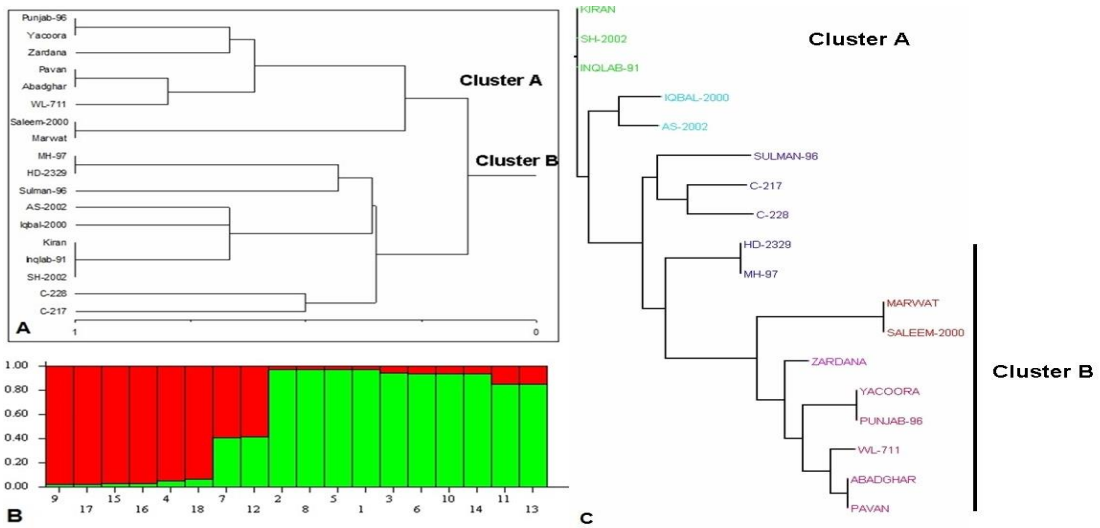

Fig. 3. Bioinformatics analysis of seed storage proteins of wheat. **A:** Phylogenetic analysis of different wheat varieties showing diversity among seed storage proteins, **B:** Grouping of 18 wheat varieties by using STRUCTURE, and **C:** Grouping of wheat varieties by TASSEL

Dendrogram revealed 2 main clusters differentiated as A and B at the distance of nearly 85%. Cluster A comprised of eight varieties, namely Punjab-96, Yacora, Zardana, Pavon, Abadgar, WL-711, Saleem-2000 and Merwat. In this cluster Pavon and Abadgar, Punjab-96 and Yacora and Saleem-2000 and Merwat showed 100% similarities. Members of cluster B were C-228, Sulman-96, C-217, Kiran, Inqilab-91, Iqbal-2000, AS-2002 and SH-2002. In cluster B 100%

similarities were shown between MH-97 and HD-2329 and among Kiran, Inqilab-91 and SH-2002. STRUCTURE was used for the analysis of population structure. This tool performed the grouping of 18 wheat varieties into two groups as shown in the Fig. 4. TASSEL was used for cluster analysis. This tool also grouped the 18 varieties into 2 main groups containing same individuals as described by other tools.

4. Discussion

The storage proteins are not affected by environmental changes, therefore protein electrophoresis is a powerful tool for population genetics and SDS – PAGE is a reliable tool for economic characterization of germplasm (Javid et al., 2004; Iqbal et al., 2005). Seeds are the most important plant storage organ and play an important role in the life cycle of plants. Since little is known about the protein composition of seed storage protein of wheat (Larre et al., 2010), their comparison has been found to provide no biological basis for unraveling closely connected small and large seeded lentils (Ladizinsky, 1979). Similar results were obtained in our genotypes investigated in the present study.

Our results demonstrated the low degree of heterogeneity in the overall form of seed storage proteins. The diversity in high molecular protein subunits is the outcome of gene silencing in some varieties encoding these proteins (Lawrence and Shepherd, 1980). With a review of the results, it is evident that polymorphism of HMW-GS existed among the Pakistani wheat varieties as described by many researchers (Chaparzadeh and Sofalian, 2007). In this study, Glu-1A locus of A genome contributed three alleles i.e, Null, 1 and 2* with the dominance of 2* subunit 55.55% and is comparable to the reports of Lee et al., (2007) and Chaparzadeh and Sofalian, (2007). The presence of Null allele was relatively low as compared to 2*, but higher than the allele 1 having lowest frequency (16.67%). The presence of subunits 1 has been reported in European wheat landraces and obsolete cultivars (Gregova et al., 2006).

Glu-1B locus showed higher levels of polymorphism by contributing 5 different types of allelic variants. Previously, Pike and MacRitchie, (2004) and Yang et al., (2007) also observed maximum glutenin polymorphism at glu-B1 locus. The most frequent pattern was 17+18 (61.11%) and are in line with the findings of Singh et al., (2007) and Tahir et al., (1995) who reported a higher frequency of subunit pair 17+18 but not in agreement with other scientists (Gregova et al., 2006; Popa et al., 2003 and Lee et al., 2007). The Glu-D1 locus of the D genome also showed a considerable amount of polymorphism. The alleles encoding the subunits pairs 5+10 and 2+12 were the most frequent patterns. The dominant proportion of 2+12 and 5+10 subunits had already been covered in various collections in different studies (Nakamura, 2001; Popa et al., 2003; Lee et al., 2007), although Valizadeh, (2001) and Chaparzadeh and Sofalian, (2007) observed the lower frequency of 5+10 subunits in Iranian genotypes.

STATISTIXL performed a cluster analysis using the similarity measure of Jaccard. The results were displayed in the form of dendrogram on the basis of group average. Dendrogram revealed 2 main clusters differentiated as A and B at the distance of about 85%. The data obtained from cluster analysis is useful to identify contrasting parents coupled with the close genetic relatedness among various crop species for better development of hybrid identification and generation of wider variability for crop improvement (Maity et al., 2004). Likewise, wheat genotypes in this field indicated no clear differentiation regarding the origin of genotypes but exhibited homology among each other.

Acknowledgments

The authors express their genuine thanks to the Nuclear Institute for Agriculture and Biology (NIAB), Faisalabad, Pakistan, for providing research facilities.

References

Chaparzadeh, N., & Sofalian, O. (2007). Study of glutenin subunits in some wheat landraces from northwest of Iran by SDS–PAGE technique. *International Journal of Agriculture and Biology*, 10(1), 101–104.

Economic Survey of Pakistan 2011–12. (2012). Ministry of Finance, Government of Pakistan.

Farooq, M. A., Iqbal, M., Masood, M. S., Rabbani, M. A., & Munir, M. (2010). Assessment of genetic diversity among Pakistani wheat (*Triticum aestivum* L.) advanced breeding lines using RAPD and SDS–PAGE. *Electronic Journal of Biotechnology*, doi: 10.2225/vol13-issue3-fulltext-2.

Gregova, E., Hermuth, J., Kraic, J., & Dotlacil, L. (2006). Protein heterogeneity in European wheat landraces and obsolete cultivars. *Genetic Resources and Crop Evolution*, 53(5), 867–71.

Huang, D. Y., & Khan, K. (1997). Quantitative Determination of High Molecular Weight Glutenin Subunits of Hard Red Spring Wheat by SDS – PAGE. I. Quantitative Effects of Total Amounts on Breadmaking Quality Characteristics. *Cereal*

Chemistry, 74(6), 781–785.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, 227(5259), 680–685.

Larre, C., Penninck, S., Bouchet, B., Lollier, V., Tranquet, O., Denery–Papini, S., Guillon, F., & Rogniaux H. (2010). Brachypodium distachyon grain: identification and subcellular localization of storage proteins. *Journal of Experimental Botany*, 61(6), 1771–1783.

Lawrence, G. J., & Shepherd, K. W. (1981). Chromosomal location of genes controlling seed proteins in species related to wheat. *Theoretical and Applied Genetics*, 59, 25–31.

Lee, E. Y., Bang, J. Y., Park, G. W., Choi, D. S., Kang, J. S., & Kim, H. J., et al. (2007). Global proteomic profiling of native outer membrane vesicles derived from *Escherichia coli*. *Proteomics*, 7(17), 3143–53.

Maity, J. P., Chakraborty, A., Saha, A., Santra, S. C., & Chanda, S. (2004). Radiation induced effects on some common storage edible seeds in India infested with surface microflora. *Radiation Physics and Chemistry*, 71, 1065–107.

Nakamura, H. (2001). Genetic diversity of high molecular weight glutenin subunit composition in landraces of hexaploid wheat from Japan. *Euphytica*,120, 227–234.

Omid, S., & Valizadeh, M. (2009). Investigation of Seed Storage Proteins in some Wild Wheat Progenitors Using SDS–PAGE and Acid–PAGE. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca*, 37(1), 179–182.

Payne, P. I., & Lawrence, C. J. (1983). Catalogue of alleles for the complex gene loci, Glu - A1, Glu - B1 and Glu - D1 which code for high molecular weight subunits of glutenin in hexaploid wheat. *Cereal Research Communications*, 11, 29–35.

Popa, M., Gregova, E., & Kraic, J. (2003). Romanian wheat (*Triticum aestivum* L.) landraces characterized by seed storage proteins. *Biodiversity FAO*, 135, 53–8.

Pike, P. R., & MacRitchie, F. (2004). Protein composition and quality of some new Hard White Winter Wheats. *Crop Science*, 44, 173–176.

Rohlf, F. J. (2000). NTSYS-pc, Numerical Taxonomy and Multivariate Analysis System, Version 2.1 Exeter Pulications, New York, USA.

Shewry, P. R., & Halford, N. G. (2002). Cereal seed storage proteins: structures, properties and role in grain utilization. *Journal of Experimental Botany*, 53(370), 947–958.

Singh, A. M., Deveshwar, J. J., Tahlawa, A. K., & Singh, B. B. (2007). Identification of novel variants of high molecular weight glutenin subunits in Indian bread wheat landraces. *Cereal Research Communications*, 35(1), 99–108.

Tahir, M., Tuchetta, H. T., & Lafiandra, D. (1995). The HMW glutenin subunit composition of bread wheat varieties bred in Pakistan. *Plant Breeding*, 114, 442–444.

Iqbal, S. H., Ghafoor, A., & Ayub, N. (2005). Relationship between SDS–PAGE markers and *Ascochyta* blight in chickpea. *Pakistan Journal of Botany*, 37, 87–96.

Javid, A., Ghafoor, A., & Anwar, R. (2004). Seed storage protein electrophoresis in groundnut for evaluating genetic diversity. *Pakistan Journal of Botany*, 36, 87–96.

Ladizinsky, G. (1979). Species relationships in the genus *Lens* as indicated by seed protein electrophoresis. *Botanical Gazette*, 140, 449–451.

Yang, P., Chen, H., Liang, Y., & Shen, S. (2007). Proteomic analysis of de-etiolated rice seedlings upon exposure to light. *Proteomics*, 7, 2459–2468.

Valizadeh, M. (2001). Seed Storage Protein Profile of Grain Legumes Grown in Iran, Using SDS–PAGE. *Journal of Agricultural Science and Technology*, 3, 287–292.

Zilic, S., Barac, M., Pesic, M., Dodig, D., & Ignjatovic-Micic, D. (2011). Characterization of Proteins from Grain of Different Bread and Durum Wheat Genotypes. *International Journal of Molecular Sciences*, 12(9), 5878–5894.