*Australian Journal of Crop Science*
*3(2):107-112 (2009)*
*ISSN: 1835-2707*

Southern Cross Journals©2009
www.cropj.com

# Evaluation of genetic diversity in soybean (*Glycine max*) lines using seed protein electrophoresis

**[*1]M. Faisal Anwar Malik, Afsari S. [1]Qureshi, [2]Muhammad Ashraf, [1]Muhammad Rashid Khan, and [2]Asif Javed**

*[*1] Department of Biochemistry, Quaid-i-Azam University, Islamabad.*
*[2] National Agricultural Research Centre, Islamabad, Pakistan.*

*****Corresponding author: malikfaisal77@hotmail.com*

**Abstract**

The genetic variation of seed protein was assayed by SDS-PAGE for ninety-two accessions of soybean (*Glycine max*). The germplasm represented five different origins/sources (Pakistan, USA, AVRDC, North Korea and Japan). To our knowledge, no studies have yet been made in Pakistan on the diversity of soybean germplasm based on protein electrophoresis. On the basis of SDS-PAGE, 26 reproducible bands were used for analysis and genetic diversity was estimated based on the number of different protein peptides. Ten major bands were recorded out of total 26 bands detected, while 50% of total were polymorphic. Dendrogram constructed using Ward's method divided the accessions in two main groups consisting of four clusters. The results of cluster analysis indicated that genetic diversity between Pakistani and US or AVRDC accessions is much larger that the genetic diversity between Pakistani and North Korean or Japanese accessions. Although cluster analysis completely separated most of the Pakistani accessions from USA and AVRDC accessions, but could not distinguish between the accessions from Japan and North Korea. As the accessions from various sources differed considerably, it was difficult to establish any relationship between origin and clustering pattern.

***Keywords:*** Genetic variation; SDS-PAGE; protein electrophoresis; cluster analysis; soybean; *Glycine max.*

## Introduction

Soybean is one of the non-conventional oilseed crops which can be successfully grown in the country such as Pakistan during both spring and the autumn seasons. It contains 40 to 42% good quality protein and 18 to 22% oil comprising 85% unsaturated fatty acids and is free from cholesterol, so it is highly desirable in the human diet (Aslam *et al*., 1995). In Pakistan it is cultivated under a wide range of agro-ecological zones particularly under rainfed conditions. At present it is cultivated over an area of 298 ha with a production of 371 tonnes. National average yield of 1250 kg/ha is very low compared with its potential, and yields obtained in other soybean producing countries (Anonymous, 2006). Soybean among food legumes are recognized as important crops all over the world. Their cultivated types and cultivars differ greatly from each other, depending on different growing regions. It is valued judgment to understand their genetic variability and relationship for facilitating the transfer of useful genes among cultivated species and maximizing the use of available germplasm resources. The extent of genetic diversity in germplasm can be assessed through morphological characterization and genetic

***Table 1.*** Variation in banding pattern of seed proteins based on origin/source in ninety-two accessions of soybean

| Sr. No. | Origin source | Accessions | No. of Bands | | Band Freq | |
|---|---|---|---|---|---|---|
| | | | Total | Polymorphic (%) | Mean | Range |
| 1 | AVRDC | 20 | 23 | 39.13 | 0.58 | 0.15-0.95 |
| 2 | Japan | 8 | 24 | 29.17 | 0.38 | 0.13-0.88 |
| 3 | N. Korea | 4 | 25 | 8.00 | 0.75 | 0.75-0.75 |
| 4 | Pakistan | 26 | 26 | 42.31 | 0.47 | 0.12-0.77 |
| 5 | USA | 30 | 25 | 48.00 | 0.55 | 0.03-0.97 |
| 6 | Checks | 4 | 23 | 26.09 | 0.42 | 0.25-0.5 |

markers. The characterized material then helps the plant breeders to select the accessions to be utilized in hybridization programme (Ghafoor *et al*., 2002). The genetic structure of the Asian soybean population, however, still remains unclear despite its usefulness as a genetic resource (Abe *et al.,* 2003).

The introduction of molecular techniques has made possible a more accurate evaluation of the genetic and environmental components of variation, bringing greater precision to measures of genetic diversity. Among biochemical techniques, DNA molecular markers, currently in use, are too expensive as compared to protein molecular markers, which are less expensive. Electrophoresis (SDS-PAGE) is widely used to describe seed protein diversity of crop germplasm (Das & Mukarjee, 1995). This method can also be used as a promising tool for distinguishing cultivars of particular crop species (Camps *et al.*, 1994 and Jha & Ohri, 1996). However, a few studies indicated that cultivar identification was not possible with the SDS-PAGE method (De-Vries, 1996). The SDS-PAGE is a practical and reliable method for species identification because seed storage proteins are largely independent of environmental fluctuation (Gepts, 1989). Genetic diversity and the pattern of variation in the Asian soybean population have been evaluated with seed protein (Han *et al*., 1999 and Hirata *et al.,* 1999). Bushehri *et al*., (2000) evaluated twenty one soybean (*Glycine max*) cultivars electrophoretically for the banding pattern of storage proteins and suggested that SDS-PAGE is a more powerful tool to characterize soybean cultivars compared to isozyme patterns. Dobhal (1995) revealed significant variability among soybean accessions for yield components, allowing accessions to be grouped into 17 clusters. There was no linear relationship between geographic and genetic distance. In view of the developing country like Pakistan, which has a great demand of soybean, such information was not available. To our knowledge, no studies have yet been made in Pakistan on the diversity of soybean germplasm based on protein electrophoresis. So the present study aimed, to evaluate the genetic variation in soybean using 92 accessions from five different origins/sources; and to relate genetic diversity patterns to geographical regions.

**Materials and methods**

Ninety-two accessions from five different origins/sources (Pakistan, USA, AVRDC, North Korea and Japan) were used for SDS-PAGE. Thirty accessions were collected from different parts of Pakistan, while rest of the accessions were plant introductions. Four local check varieties (NARC-I, NARC-III, NARC-IV and NARC-V) were also included. For the extraction of proteins for electrophoresis, seeds were ground to fine powder with mortar and pestle. Sample buffer (400µl) was added to 0.01g of seed flour as extraction liquid and Bromophenol Blue (BPB) as tracking dye to follow the movement of protein in the gel. The proteins were extracted with the extraction buffer containing the following final concentrations: 0.5 M Tris-Hcl (pH 6.8), 2.5% SDS, 10% glycerol and 5% 2-mercaptoethanol. Seed proteins were separated by carrying out electrophoresis in the discontinuous buffer system through vertical slab type SDS-PAGE using 11.25% polyacrylamide gel according to the method of Laemmli, (1 970). The gels were later stained in 0.04% Coomassie Brilliant Blue in methanol, acetic acid and distilled water (45:10:45 v/v) for 8-10 hours followed by destaining in the same solution without Coomassie Brilliant Blue with occasional shaking till the gels became clear.

***Table 2.*** Clusters based on SDS-PAGE in ninety-two accessions of soybean (*Glycine max*)

| Group | Cluster | Frequency | Accessions |
|---|---|---|---|
| Group A | Cluster I | 31 | US03905, US03904, PK03903, US03901, US03900, PK03899, US03898, US03896, AV038584, AV038583, AV038581, AV03858, AV03857, AV03854, US03853, US03851, US038501, US03850, US038491, US03849, US03848, US03893, AV038585, AV03855, US038521, AV03840, US03838, AV03839, US03894, NARCIV, JA03832. |
| | Cluster II | 23 | AV03844, AV03842, AV038401, US03845, JA03830, AV03843, AV03841, JA03828, US03907, US03906, US03897, US03902, AV038582, US03852, AV03856, NARCV, JA03834, US03895, US03847, US03846, JA03837, JA03836, JA03831. |
| Group B | Cluster III | 21 | PK03777, PK03776, PK03774, PK03773, NARCIII, NARCI, AV038586, US038591, AV038588, AV038587, US038590, US038589, US038592, JA03826, NK03792, NK03818, NK03817, NK03816, PK037821, PK03782, PK037721. |
| | Cluster IV | 17 | PK037621, PK03762, PK03758, PK037722, PK037671, PK03768, PK03763, PK03757, PK037551, PK03769, PK03755, PK03767, PK037661, PK03766, PK03765, PK03772, PK03754. |

The resulting gels were photographed to visualize the protein band patterns.

**Data analysis**

After staining and destaining the gels, the number of monomorphic and polymorphic protein bands were counted for each sample. The presence (1) or absence (0) of polypeptide bands was entered in a binary data matrix for use in cluster analysis. The polymorphic bands were analysed for the level of polymorphism by counting the number of polymorphic bands and generating summary statistics on the band frequencies. Cluster analysis was also performed using the computer software Statistica.

**Results and discussion**

On the basis of the relative mobility of seed proteins on the gel, 26 bands were detected in this study, which were used for examining the genetic diversity (Fig 1). Ten major bands were recorded out of total 26 bands detected, while 50% of total were polymorphic. The banding pattern revealed large variations among accessions. The banding pattern revealed three regions. Region I comprised of 11 bands of above 66.0 kD molecular weight (MW), Region II consisted of nine bands lying between 45.0 and 66 kD, while Region III had six bands between 24.0 kD and 45.0 kD. The bands below 24.0 kD were diffused and not considered for analysis purpose.

The band frequencies in these samples ranged from 0.03 to 0.99 with a mean of 0.53. The results of Table 1 showed that the total number of seed storage protein bands observed for each origin/source varied from 23 to 26 and averaged 24.3. The percentages of polymorphic bands over the total bands detected ranged from 8% (N. Korea) to 48% (USA) and averaged 32%. The mean band frequency for each population ranged from 0.38 (Japan) to 0.75 (Japan) with a mean value of 0.53.

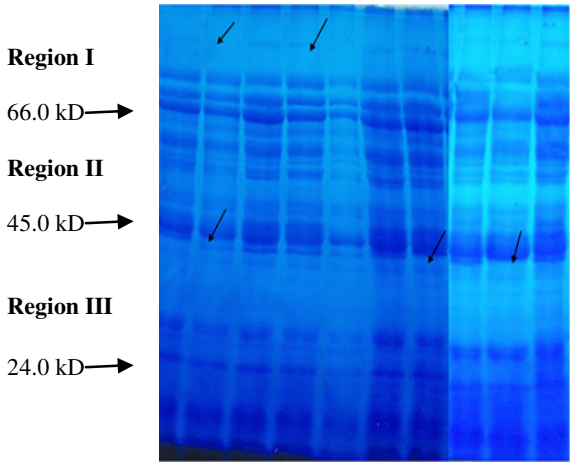


***Fig 1.*** Banding pattern of SDS-PAGE showing diversity among soybean accessions. The arrows indicate the diversity in bands.

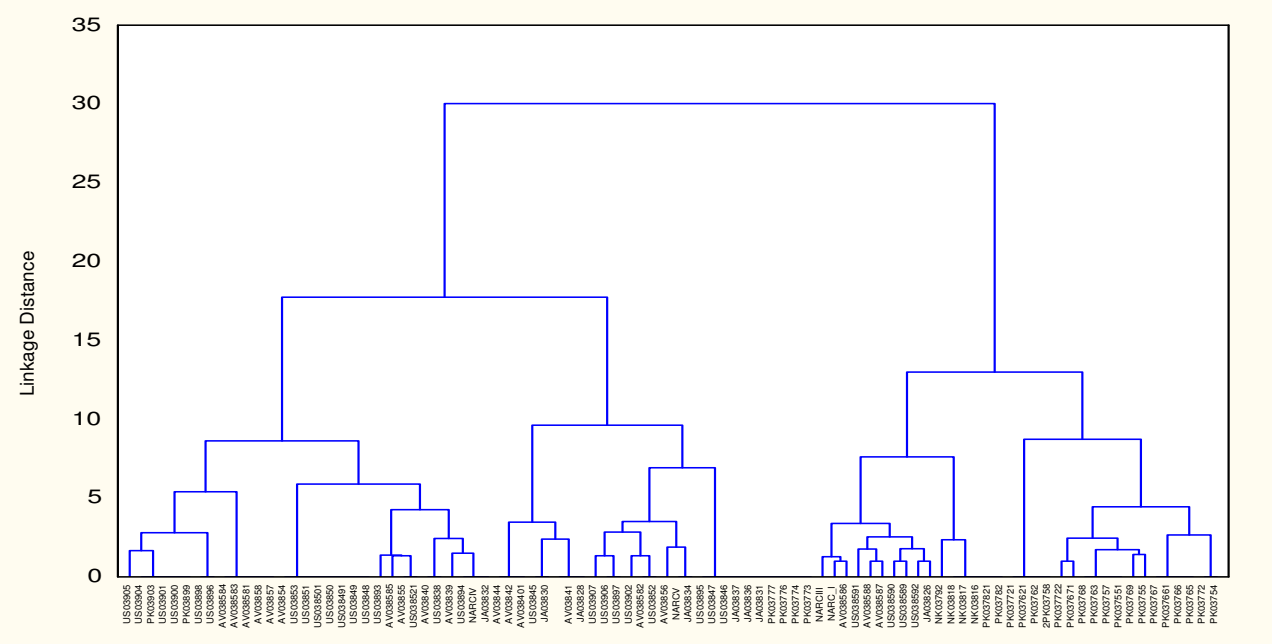


***Fig 2.*** Dandrogram based on SDS-PAGE of ninety-two accessions of soybean by Ward's method.

The percentage of polymorphic bands observed for accessions of each origin/source was associated with the number of samples assayed. Mirza *et al.,* (2007) observed that the total number of seed storage protein bands for each population of *Avena fatua* varied from 26 to 34 and averaged 31.

The percentages of polymorphic bands over the total bands detected ranged from 37.9 (Lakkimarwat population) to 97% (Swabi population) and averaged 73.4%. The mean band frequency for each population ranged from 0.41 (Islamabad) to 0.66 (Gujranwala; Lakkimarwat and Mardan) and averaged 0.55. The molecular weights for these bands ranged from 14.2 to 66 kD.

Dendrogram of total seed protein using Ward's method showed that the accessions were divided into two main groups consisting of four clusters (Fig. 2). Cluster I consisted of 17 accessions of US (US03905, US03904, US03901, US03900, US03898, US03896, US03853, US03851, US038501, US03850, US03-8491, US03849, US03848, US03893, US038521, US03838 and US03894), 10 of AVRDC (AV038584, AV038583, AV038581, AV03858, AV03857, AV03854, AV038585, AV03855, AV03840, and AV03839), 2 of Pakistani (PK03903 and PK03899) and 1 of Japanese origin (JA03832). Except two Pakistani accessions, which included in cluster I, all the Pakistani accessions were included in cluster III (PK03777, PK03776, PK03774, PK03773, PK037-821, PK03782, PK037721 and cluster IV (PK037621, PK03762, PK03758, PK037722, PK037671, PK03-768, PK03763, PK03757, PK037551, PK03769, PK03755, PK03767, PK037661, PK03766, PK03765, PK03772, PK03754). It is clear that cluster IV consisted of only Pakistani accessions. The checks were scattered in cluster I (NARC IV), cluster II (NARC V), and cluster III (NARC III and NARC I). All the North Korean accessions (NK03792, NK03818, NK03817, NK03816) were grouped into cluster III, while majority of the American and AVRDC accessions were included in cluster I and II (Table 2). Bushehri *et al*., (2000) also evaluated twenty-one soybean (*Glycine max*) cultivars electrophoretically using polyacrylamide and starch gel electrophoresis and found that the banding patterns of seed storage proteins classified cultivars into 4 distinct clusters.

The results of cluster analysis indicated that genetic diversity between Pakistani and US or AVRDC accessions is much larger than that the genetic diversity between Pakistani and North Korean or Japanese accessions. Although cluster analysis completely separated most of the Pakistani accessions from USA and AVRDC accessions, but could not

distinguish between the accessions from Japan and North Korea. On the basis of these results, it is clear that crosses between the Pakistani and US or AVRDC gene pools could create more genetic variability than crosses between Pakistani and Japanese or North Korean accessions. As the accessions from various sources differed considerably, it was difficult to establish any relationship between origin and clustering pattern. Sihag *et al*., (2004) found that the cluster pattern obtained showed that genetic diversity and geographic distribution were independent of each other and no definite relationship existed between genetic diversity and geographic diversity. Alipour *et al*., (2002) conducted an experiment to study the genetic variation in the electrophoretic patterns of seed proteins. Based on the relative mobility on the gel, 30 protein bands were observed, of which only 5 bands varied among the accessions. Cluster analysis based on qualitative evaluation of the patterns grouped the accessions into 8 clusters and classified the different bands into 3 groups. Ghafoor *et al.,* (2003) also reported similar results. SDS-PAGE cannot be used for identification of various genotypes on the basis of intraspecific variation, because some of the accessions that differed on the basis of characterization and evaluation exhibited similar banding patterns; this technique might be used to study inter rather than intraspecific variation in Vigna spp. (Ghafoor *et al*., 2002).

Results are somewhat contradictory to the findings of Javaid *et al.,* (2004) who reported low genetic diversity in groundnut for SDS-PAGE, and suggested 2-D electrophoresis. Genotypes with similar banding patterns are suggested to be studied for detailed agronomic and biochemical analyses, including 2-D electrophoresis and DNA markers, for better management of the genebank (Celis and Bravo, 1984; Beckstrom-Sternberg, 1989).

It was concluded that the hybridization between Pakistani and US or AVRDC accessions would be more useful as they showed greater genetic diversity than the genetic diversity between Pakistani and N. Korean or Japanese accessions.

**References**


Abe J, Xu DH, Suzuki Y, Kanazawa A, Shimamoto Y (2003) Soybean germplasm pools in Asia revealed by nuclear SSRs. Theor. Appl. Genet. 106:445–453.

Alipour H, Rezai A, Meibodi SAM, Taheri M (2002) Evaluation of genetic variation in soybean lines using seed protein electrophoresis. J. Sci. Technol. Agric. Nat. Resour. 5(4): 85-96.

Anonymous (2006) Agri. Statistics Pak. Pg 65.

Aslam M, Mirza MS, Shah SM, Shafeeq S, Naeem Ullah (1995). Crop Production Bulletin No. 6, Pak. Agric. Res. Council, Islamabad, Pakistan. pg. 1.

Beckstrom-Sternberg SM (1989) Two-dimensional gel electrophoresis as a taxonomic tool: evidence from the Centrospermae. Biochem. Syst Ecol. 17: 573–582.

Bushehri SAA, Abd-Mishani C, Yazdi-Samadi B, Sayed-Tabatabaei BE (2000) Variety specific electrophoretic profiles of soybean cultivars. Iranian J. Agric. Sci. 31(1): 55-61.

Camps G, Vernetti Fde-J, Augustin E, Irigon D (1994) Morphological and electrophoretical characterization of twenty soybean cultivars. Pesqui. Agropecu. Bras. 29(11): 1779-1787.

Celis JE, Bravo R (1984) Two dimensional gel electrophoresis of proteins; methods and applications. Acad. Press, New York.

Das S, Mukarjee KK (1995) Comparative study on seed proteins of *Ipomea*. Seed Sci. Technol. 23: 501-509.

De-Vries, IM (1996) Characterization and identification of *Lactuca sativa* cultivars and wild relatives with SDS-electrophoresis (*Lactuca* sect. *Lactuca*, *Compositae*) GRACE. 43: 193-202.

Dobhal VK (1995) Genetic divergence in soybean. Legume Res. 18(1): 29-34.

Gepts P (1989) Genetic diversity of seed storage proteins in plants. In: A.H.D. Brown, M.T. Clegg, A.L. Kahler & B.S. Weir (Eds.), Plant Population Genetics, Breeding and Genetic Resources. pp. 64-82. Sinauer Associates Inc., Sunderland, Massachusetts.

Ghafoor A, Zahoor A, Qureshi AS, Bashir M (2002) Genetic relationship in *Vigna mungo* (L.) Hepper and *V. radiata* (L.) R. Wilczek based on morphological traits and SDS-PAGE. Euphytica. 123: 367-2002.

Ghafoor A, Gulbaaz FN, Afzal M, Ashraf M, Arshad M (2003). Inter-relationship between SDS-PAGE markers and agronomic traits in Chickpea (*Cicer Arietinum*). Pak. J. Bot. 35(4): 613-624.

Han O, Abe J, Shimamoto Y (1999) Genetic diversity of soybean landraces in Korea. Korean J. Crop Sci. 44: 256–262.

Hirata T, Abe J, Shimamoto Y (1999) Genetic structure of the Japanese soybean population. Genet. Resour. Crop Evol. 46: 441–453.

Javaid, A, Ghafoor A, Anwar R (2004). Seed storage protein electrophoresis in groundnut for evaluating genetic diversity. Pak. J. Bot. 30(1): 25-29.

Jha SS, Ohri D (1996) Phylogenetic relationships of *Cajanus cajan* (L.) Millsp. (pigeonpea) and its wild relatives based on seed protein profiles. Genet. Resour. Crop Evol. 43: 275–281.

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage $T_4$. Nature. 227: 680-685.

Mirza B, Shoaib M, Ahmad M, Fu YB (2007) Genetic diversity in Pakistani populations of *Avena fatua* revealed by seed storage protein polymorphism. Commun. Biometry Crop Sci. 2 (1), 41–48.

Sihag R, Hooda JS, Vashishtha RD, Malik BPS. 2004. Genetic divergence in soybean [*Glycine max* (L.) Merrill]. Annals Biol. 20(1): 17-21