Genetic Diversity and Phylogenetic Relationship Among Soybean [*Glycine max* (L.) Merrill] Varieties Based on Protein, Evolutionary and Morphological Markers

SC Sharma[1] and SR Maloo[2]

[1] *Assistant Research Scientist, Main Pulses Research Station, S.D. Agricultural University, Sardarkrushinagar-385506, (BK), Gujarat, India*

[2] *Department of Plant Breeding and Genetics, Rajasthan College of Agriculture, Maharana Pratap University of Agriculture and Technology, Udaipur-313001, Rajasthan, India*

Preliminary biochemical study was carried out to identify 14 soybean [*Glycine max* (L.) Merrill] varieties on the basis of electrophoregrams of soluble seed proteins by Discontinuous Polyacrylamide Slab Gel Electrophoresis method. Genetic diversity among soybean varieties was determined on the basis of their evolutionary, morphological and biochemical markers. Relationship was also established between varietal characters based on aforesaid parameters. In the electrophoretic study clear band patterns were observed for all 14 varieties. All type of protein band intensity i.e. low, medium and high were observed in various soybean varieties. However, in some varieties the quantitatively similar number of total protein bands were noted, but differences in presence and/or absence of particular band at particular position and their Rm values as well as different protein band intensity for common bands showed diverse nature of these varieties to each other. Results revealed that 12 varieties could easily be identified from each other as they showed their own specific characters. The variety PK 327 and JS 79-81 showed genetic diversity for their evolutionary and morphological markers, but at biochemical level they depicted identical protein banding patterns having variation only in band intensity. Such a set of varieties could possibly be discriminated by performing Slab Gel Electrophoresis under denaturing conditions or by PCR-based DNA fingerprinting. These findings could open a scope for further research in the specific area of "Varietal identification". Therefore it is concluded that all 14 soybean varieties studied were genetically diverse and could be used in breeding programme. Hence, the DPAGE technique can be effectively used for varietal identification which is found to be quick, reliable, economical, avoiding field studies.

Key Words: Variety identification, Electrophoresis, Markers, Soybean

Introduction

Variety identification with respect to its genetic purity is important in national and international seed and breeding programmes. Different cultivars are commonly identified on the basis of taxonomic differences of seed, seedling and mature plants. Distinguishing varieties on the basis of morphological characters of plants in field and seed morphology is not always possible, though it is undoubtedly, one of the most commonly used criterion. Further, parameters are highly susceptible to environmental, physiological and ecological conditions. Therefore, precise characterization of varieties should be done at both levels, viz. morphological as well as biochemical levels (McDonald, 1997). Proteins being the direct gene products, reflect the genomic composition of varieties accurately and therefore, are ideal for genotypic distinctness. Biochemical approaches like gel electrophoresis of proteins and isoenzymes are powerful tools to distinguish varieties. Seed proteins are known to controlled by multigene families (Lumen, 1990). Deletion or mutation in these structural genes or their regulatory loci results in the inhibition of transcription or translation of polypeptides and may lead to failure of protein expression (Brown *et al.*, 1981). Expression of these proteins is governed monogenically, presence being completely dominant over absence. Polypeptides varying for presence or absence could be used as markers (Naik, 1998). In the present investigation genetic diversity among soybean varieties was determined on the basis of their evolutionary, morphological and protein markers. Attempts were also made to establish the relationship between varietal characters on the basis of morphological characters and biochemical profile analysed through Discontinuous Polyacrylamide Slab Gel Electrophoresis (DPAGE) method by Davis (1964).

Materials and Methods

Healthy seeds of 14 soybean varieties namely JS 335, MACS 57, Monetta, NRC12, PK 416, PK 564, Pusa 20, Pusa 24, PK 327, JS 79-81, RAUS 97-1, Pusa 22, Pusa 20 and PK 471 were procured from soybean breeder and field experiment was conducted at Instructional Farm,

Rajasthan College of Agriculture, Maharana Pratap University of Agriculture and Technology, Udaipur. Each variety was planted in one row of 3 m length in 3 replications and seeds were dibbled at the spacing of 45 cm x 10 cm row to row and plant to plant. Observations for quantitative characters were recorded on 10 randomly selected competitive plants. Seed protein content was estimated as per standard micro–Kjeldhal method.

The experiment for identification of soybean varieties on the basis of electrophoregrams of soluble seed protein was carried out at Biochemistry and Plant Physiology Laboratory, Seed Technology Research, Rajasthan Agricultural University, Agricultural Research Station, Durgapura, Jaipur. Seeds of each variety were crushed and powdered using pestle and mortar. Seed powder (0.5 g) was vigorously mixed in ethylacetate (5 ml) for 20 minutes with the simultaneous addition of sodium sulphite (1 mg) and sodium metabisulphite (1 mg) and centrifuged at 5000 RPM for 10 minutes. The ethylacetate is decanted off and residue was treated with CMA solution (Chloroform : Methanol: Acetone, 2:1:1, v/v/v, respectively) for 20 minutes with vortexing and then solution was decanted off. The treatment was repeated one more time as above and centrifuged at 5000 RPM for 10 minutes. The final residue was kept in suitable volume (1 ml) of extraction buffer (0.1 M Tris-HCl buffer pH 7.5) for 18 hours in cold. Contents were then centrifuged at 10,000 RPM for 15 minutes. The clear supernatant was used as protein sample for electrophoretic analysis. During protein extraction, treatments with ethylacetate and CMA solution removed most of the lipids, as presence of lipids interferes during isolation of proteins (Stegemann and Pietsch, 1983). To prevent auto-oxidation of phenols, sodium sulphite and sodium metabisulphite were added. The loss of proteins by above treatment was negligible. Discontinuous Polyacrylamide Slab Gel Elctrophoresis (DPAGE) was carried out according to standard method (Davis, 1964) using 12% acrylamide separating gel (0.5 M Tris-HCl buffer pH 8.8) with a top layer of 6% acrylamide stacking gel (0.5 M Tris-HCl buffer pH 6.8). The standardized volume of protein sample (40 µl) was loded on slab gel wells. Electrophoresis was conducted at a constant current of 36 mA for 45 minutes and then raised to 58 mA until the tracking dye (bromophenol blue) migrated to the anode end of the slab gel. The staining was done for 2 hours in coomassie brilliant blue R-250 (CBBR) solution (0.5 g CBBR dye was dissolved in 250 ml methanol, 10 ml acetic acid and 240 ml water). The gels were destained by repeated washing with methanol : acetic acid : water (50:70:880, v/v/v, respectively). The protein bands were numbered from cathode end and their relative mobilities (Rm) were calculated.

Results and Discussion

Analysis of seed proteins using electrophoresis (PAGE) has been employed for identification of varieties and elucidation of evaluation and phylogenetic relationship in soybean viz. Hilty and Schmitthenner (1966), Larson (1967), Wagner and McDonald (1982), Cardy and Bevers darg (1984), Kumar and Ram (1989), McDonald (1982), Goyal and Sharma (1988) and Anonymous (2003).

The protein banding pattern of all 14 soybean varieties were compared and are shown in Figure 1 and 2 and

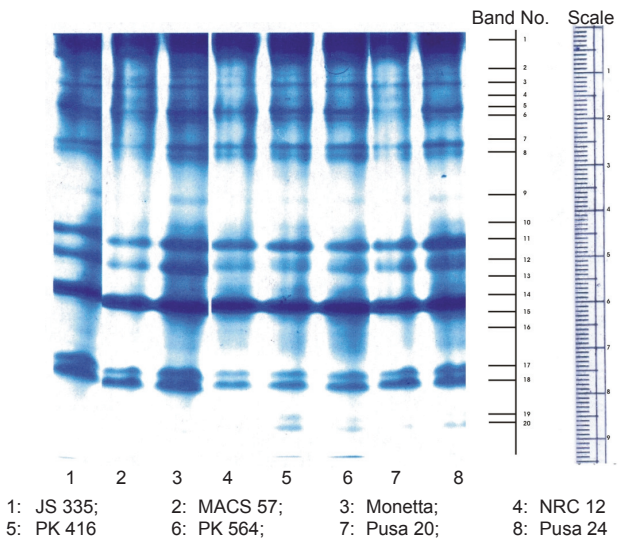

Fig. 1: Protein banding patterns of soybean varieties

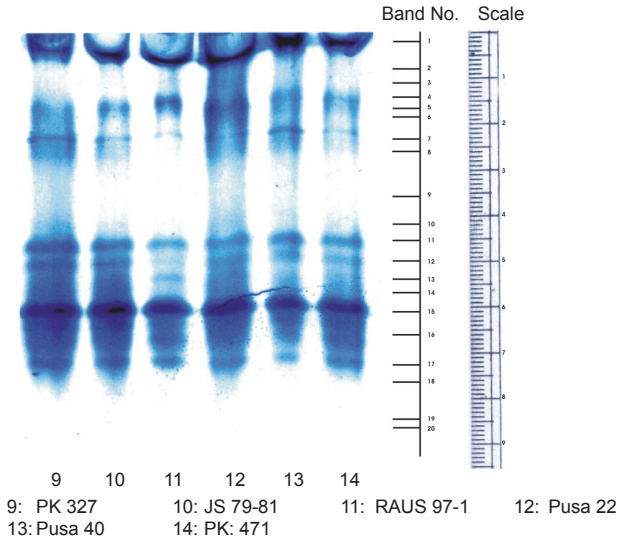

Fig. 2: Protein banding patterns of soybean varieties

results are presented in Table 1 and 2. The inter-varietal differences were clearly expressed in these patterns. Rm values ranged from 0.021 to 0.920 and indicated different mobility pattern thereby suggesting wider range of variability in protein band expression. In all 20 protein bands of different Rm values were identified on the basis of electrophoretic mobility among 14 soybean varieties being maximum (17) in PK 416 and minimum (7) in PK 327 and JS 79-81 (Table 2). Sampat (1981) also reported 18-25 protein bands in different Vigna species, while total 15 protein bands in 8 soybean [*Glycine max* (L.) Merril] varieties were also observed by Goyal and Sharma (1998). The similarity in certain protein bands among varieties as observed from the Table 2 were noted and this might be due to genetic relationship among them as also reported by Goyal *et al.* (1998) in cotton (*Gossypium herbaceum*), Singh *et al.* (2002) in *Catharanthus rosesus* and Goyal and Sharma (2003) in cluster bean. On the basis of protein electrophoregram pattern, it was observed that protein band No. 2, 3, 4, 5, 6, 8, 9, 10, 13, 14, 15 and 18 were useful for discriminating varieties as they showed variation in their expression. This indicated distinct variations among the soybean varieties at biochemical level used in the study, further substantiated the evolutionary and morphological diversity (Table 1). The presence and absence of these bands could help in determination of inter-varietal differences as reported by Chand and Kole (2002) in mungbean (*Vigna radiata*).

On the other hand protein band No. 16 was distinctly present in only 2 varieties, namely, RAUS 97-1 and PK 471, while band No. 10 was found in Monetta, PK 416 and Pusa 24 (Table 2). The specific presence of these 2 bands in aforesaid varieties showed their specific identity which also reflected from their evolutionary as well as morphological markers (Table 1). It was noticed that protein band No.1, 7, 11, 12 and 17 were observed to be universally present in all 14 varieties studied indicated the genes controlling the expression of these protein bands appeared to behave a single block as also revealed by Padmavathi *et al*. (1999) in rice (*Oryza sativa*) . The protein band No. 19 and 20 were only present in 3 varieties, viz., PK 416, PK 564 and Pusa 24. For PK 416 and PK 564 this might be because of their common geographical origin and over and above on account of 2 common parents in their pedigree. However, the protein band intensity of these 2 bands was categorized as low and these bands were missing in rest all 11 varieties. Possibly, the genes controlling the expression of these bands might be linked. This assumption could, however, be verified through linkage analysis using segregating populations as also reported by Padmavathi *et al.* (1999) in rice (*Oryza sativa*).

Multigene inheritance of seed protein expression could be identified from electrophoregram. Multigene inheritance of seed protein expression is well established in other crops (Lumen, 1990). Seed protein expression is known to be under monogenic control with co-dominance of alleles of diverse molecular weight variants and dominance of the presence over absence. The association of protein bands with agrobotanic characteristics has previously been reported. Whereas Larson (1967) observed genotypes with a particular banding pattern to have linkage with hilum colour in soybean.

Polyacrylamide gel electrophoresis in different soybean varieties belonging to same geographical origin revealed variation in number, width, relative mobility and intensity of bands. In general, geographical origin does not influence genotype specific seed protein content and its polymorphism as also reported in grass pea (*Lathyrus sativus*) by Roy *et al*. (2001). The study by and large demonstrated genotype – specific protein polymorphism which is not always influenced by geographical origin.

All type of protein band intensity i.e. low, medium and high were observed in various soybean varieties. Most of the soybean varieties expressed their own specific characteristics which differentiate them from other varieties. However, in some varieties the quantitatively similar number of total protein bands were noted, but differences in presence and/or absence of particular band at particular position and their Rm values, as well as different protein band intensity for common bands showed diverse nature of these varieties to each other. Table 2 revealed total 12 protein bands in electrophoregram of variety JS 335 and NRC 12, but band position was found variable. The band No.6 was characteristically present in JS 335 while missing in NRC 12. Similarly band No. 13 was present in NRC 12 and missing in JS 335. Besides this, JS 335 and NRC 12 could be differentiated to each other considerably based on their geographical origin, evolutionary pedigree and morphological traits particularly with respect to growth habit, type, leaf intensity of green colour, seed shape, seed hilum colour and other quantitative characters studied (Table 1).

Electrophoregrams of variety MACS 57, Pusa 22 and PK 471 (Table 2) showed quantitatively equal 11

Table 1. Genetic diversity for evolutionary and morphological markers in different varieties of soybean

Characters	JS 335	MACS 57	Monetta	NRC 12	PK 416	PK 564	Pusa 20	Pusa 24	PK 327	JS7 9-81	RAUS 97-1	Pusa 22	Pusa 40	PK 471
(A) Evolutionary Markers														
Year/ Notification year	1994	1992	1985	1997	1986	1991	1988	1987	1982	1990	1989	1983	1981	1988
Geograp-hical origin	Jabalpur (JNKVV)	Pune (MACS)	Selection from exotic culture	Indore (NRCS)	Pantnagar (GBPUAT)	Pantnagar (GBPUAT)	New Delhi (IARI)	New Delhi (IARI)	Pantnagar (GBPUAT)	Jabalpur (JNKVV)	Kota (ARS,RAU)	New Delhi (IARI)	New Delhi (IARI)	Pantnagar (GBPUAT)
Evolu-tionary pedigree	JS78-77 x JS 71-05	JS2 x Improved/ Pelicon	PL71608 from Nanking, China	Mutant of Bragg	UPSM 534 x S 38	(UPSM 534 x S-38) x Bragg	Bragg x Lee	Shelby x Bragg	UPSM 82 x Sammens	Bragg x Hara Soya	Punjab 1 x Monetta	Punjab Clark 63	S3 x Lee	Hardee x Punjab 1
Area of adaptation	Central Zone	Mahara-shtra	Central and Southern Zone	Central Zone	Northern Plain and Hill Zone	Northern Plain Zone	Northern Hill Zone	Northern Plain and Central Zone	Northern Plain and Hill Zone	Central Zone	Central Zone	Northern Plain and Central Hill Zone	Southern Zone	Southern Zone
(B) MORPHOLOGICAL MARKERS														
(I) **Qualitative Characters**														
Growth habit	Semi-determinate	Semi-determinate	Determi-nate	Determi-nate	Determi-nate	Determi-nate	Semi-determinate	Semi-determinate	Determi-nate	Determi-nate	Determi-nate	Determi-nate	Determi-nate	Determi-nate
Growth type	Semi-erect	Semi-erect	Erect	Erect	Erect	Semi-erect	Semi-erect	Semi-erect	Semi-erect	Erect	Erect	Erect	Erect	Erect
Flower colour	Violet	Violet	Violet	Violet	White	White	White	White	Violet	Violet	Violet	Violet	Violet	White
Leaf intensity of green colour	Dark	Medium	Medium	Medium	Medium	Medium	Medium	Medium	Medium	Medium	Dark	Medium	Medium	Medium
Seed shape	Spherical flattened	Spherical	Spherical	Spherical	Elongated	Spherical flattened	Spherical	Spherical	Spherical	Spherical	Spherical flattened	Round	Round	Spherical
Seed hilum colour	Black	Brown	Brown	Brown	Black	Brown	Brown	Black	Brown	Brown	Brown	Black	Black	Brown
(II) **Quantitative Characters**														
Days to flowering	38.33	42.00	43.33	39.67	43.67	39.33	44.67	45.33	41.00	44.00	44.33	41.33	46.67	45.33
Days to maturity	96.67	102.00	99.67	97.00	102.33	96.33	101.67	102.00	97.00	101.33	103.33	103.33	106.00	103.67
Plant height (cm)	48.98	48.17	47.67	39.35	52.10	43.85	41.74	51.03	46.25	42.85	45.70	47.10	41.18	40.15
Pod Length (cm)	3.63	3.44	3.48	3.19	3.68	3.70	3.46	3.51	3.55	3.34	3.25	3.58	3.23	3.38
Height of pod insertion (cm)	6.32	6.05	6.88	7.56	6.15	6.05	6.43	7.13	5.64	7.29	6.95	5.71	7.44	6.10
100-seed weight (g)	10.44	9.50	9.61	8.01	10.32	9.50	10.64	11.26	11.83	10.13	9.20	11.95	9.43	8.80
Protein content (%)	39.67	39.49	38.61	38.74	39.22	39.88	37.10	39.10	39.35	37.19	37.83	38.51	36.61	39.78

Table 2. Protein banding patterns of soybean varieties

Protein band No.	Rm value	Varieties (Presence/absence of protein bands and their intensity)													
		JS 335 (12)	MACS 57 (11)	Monetta (15)	NRC 12 (12)	PK 416 (17)	PK 564 (16)	Pusa 20 (13)	Pusa 24 (16)	PK 327 (7)	JS 79–81 (7)	RAUS 97–1 (9)	Pusa 22 (11)	Pusa 40 (8)	PK 471 (11)
1.	0.021	+++	+++	+++	+++	+++	+++	+++	+++	+++	++	++	+++	+++	++
2.	0.085	++	+	++	+	++	++	+	++	–	–	–	+	–	–
3.	0.122	++	++	++	+	++	++	+	++	–	–	–	+	–	–
4.	0.146	–	+	+	–	–	–	+	–	+	++	+++	+++	+	+
5.	0.175	++	++	++	++	++	++	+	++	–	–	–	++	+	+
6.	0.191	+	–	+	–	++	++	–	++	–	–	–	–	–	–
7.	0.252	++	+	++	++	++	+++	+	+++	+++	++	+	+	++	+
8.	0.284	++	–	+++	++	++	+++	+	++	–	–	–	–	–	+
9.	0.382	–	–	++	–	+	+	+	+	–	–	–	–	–	–
10.	0.452	–	–	+	–	+	–	–	+	–	–	–	–	–	–
11.	0.486	+++	+++	+++	+++	+++	+++	+++	+++	+++	+++	++	+++	+++	+++
12.	0.537	+++	+++	+++	+++	+++	+++	+++	+++	+++	+++	+	+++	++	++
13.	0.582	–	–	–	+	+	+	–	–	–	–	+	+	–	+
14.	0.617	+++	+++	+++	+++	+++	+++	+++	+++	–	–	–	–	–	–
15.	0.654	–	–	–	–	–	–	–	–	+++	+++	+++	+++	+++	+++
16.	0.696	–	–	–	–	–	–	–	–	–	–	+	–	–	+
17.	0.781	++	++	++	++	++	++	++	++	+++	+++	+++	+++	+++	+++
18.	0.816	++	++	++	++	++	++	++	++	–	–	–	–	–	–
19.	0.898	–	–	–	–	+	+	–	+	–	–	–	–	–	–
20.	0.920	–	–	–	–	+	+	–	+	–	–	–	–	–	–

+++ : High, ++ : Medium and + : Low intensity. The values in parenthesis indicate the total number of protein bands

protein bands, but diverse banding pattern revealed clear-cut differences among them. For example, band No. 14 and 18 were present in MACS 57 and absent in Pusa 22, whereas band No. 13 present in Pusa 22 and missing in MACS 57. Similarly, presence of band No. 2, 3, 14 and 18 in MACS 57 and band No. 8, 13 and 16 in PK 471 made great variation between them. This genetic diversity was also mirrored in evolutionary and morphological features (Table1) of variety MACS 57 and Pusa 22 (geographical origin, evolutionary pedigree, growth habit, growth type, seed shape, seed hilum colour and protein content) as well as for MACS 57 and PK 471 (geographical origin, evolutionary pedigree, growth habit, growth type, flower colour and other quantitative traits). Total number of protein bands of PK564 were equivalent to that of Pusa 24 but difference in the presence and/or absence of band NO. 10 in Pusa 24 and 13 in PK 564 discriminated them (Table 2). However, variation in intensity of protein bands were also noted for different common bands. This might serve as a representative of evolutionary and morphological diversity in these 2 varieties viz., geographical origin, evolutionary pedigree, growth habit, seed shape, seed hilum colour and other quantitative traits (Table 1).

On the other hand, soybean varieties viz., PK 327 and JS 79-81 displayed a different kind of situation as they had identical banding pattern (Table 2) in their electrophoregrams indicating their high order of closeness in seed protein polymorphism/profile. However, differences were observed in the protein band intensity e.g., band No. 1 and 7 were more intense in PK 327, while band No. 4 was relatively less intense as compared to that of JS 79-81. Such type of close relationships and/or minor variations were also observed by Sasek *et al.* (1995) in barley (*Hordium vulgare*), Kharkwal (1999) in chick pea (*Cicer arietinum*), Roy *et al.* (2001) in grass pea (*Lathyrus sativus*) and Singh *et al.* (2002) in *Cathranthus rosesus*. However, evolutionary and morphological markers indicated great variability in these 2 varieties for geographical origin, evolutionary pedigree, area of adaptation, growth type and other quantitatively governed traits like days to flowering, days to maturity, plant height, 100-seed weight, pod length, height of pod insertion and protein content, hence considered as genetically diverse. But, for clear discrimination between PK 327 and JS 79-81 work on Slab Gel Electrophoresis under denaturing conditions or by PCR based DNA fingerprinting is needed.

These findings could open a scope for further research in the specific area of "Varietal identification". Therefore it is concluded that all 14 soybean varieties studied were genetically diverse and could be used in breeding programme. Hence, the DPAGE technique can be effectively used for varietal identification which is found to be quick, reliable, economical and avoiding the field work.

References

Anonymous (2003) *Morphological, chemical and electrophoretic descriptors for soybean* [*Glycine max* (L.) Merrill]. *Technical Bulletin No.15*, National Seed Project (Crops), IARI, New Delhi.

Brown JBS, TC Osborn, FA Bliss and TC Hall (1981) Genetic variation in the subunits of globulin-2 and albumin seed proteins of French bean. *Theor. Appl. Genet.* **60:** 245-250.

Cardy BJ and WD Beversdorf (1984) Identification of soybean cultivars using isoenzyme electrophoresis. *Seed Sci. & Technol.* **12:** 943-955.

Chand S and C Kole (2002) Protein markers for screening of CLS resistant mungbean [*Vigna radiata* (L.) Wilczek.] genotypes. *Indian J. Genet.* **62:** 340-341.

Davis BJ (1964) Disc electrophoresis-II. Method of application to human serum proteins. In: Gel Electrophoresis. *Ann. N.Y. Acad. Sci.* **121:** 404-427.

Goyal KC and SN Sharma (1998) Biochemical tool (Discontinuous PAGE) for varietal identification in soybean [*Glycine max* (L.) Merrill.]. *Seed Res.* **26:** 204-206.

Goyal KC and SN Sharma (2003) Biochemical approach for identification of clusterbean varieties. *Indian J. Plant Physiol.* **8:** 402-404.

Goyal KC, G Singh and SN Chaturvedi (1986) Protein electrophoregram : An identification tool for cotton seed varieties. *Agri. Biol. Res.* **2:** 17-20.

Hiltyjames W and AF Schmitthenner (1966) Electrophoretic comparison of soybean leaf proteins from varieties resistant and susceptible to *Phytophthora megasperma* var. *sojae. Phytopathol.* **56:** 287-291.

Kharkwal MC (1999) Seed storage proteins and interspecific relationship in *Cicer arietinum* L. *Indian J. Genet.* 59:59-64.

Kumar H and HH Ram (1989) Identification of soybean cultivars using electrophoretic patterns of seed proteins. *Crop Improv.* **16:** 5-8.

Larson AL (1967) Electrophoretic differences in seed proteins among varieties of soybean, *Glycine max* (L.) Merrill. *Crop Sci.* **7:** 311-313.

Lumen DEBO (1990) Molecular approaches to improving the nutritional and functional properties of plant seeds as food sources : Developments and comments. *J. Agri. Food Chem.* **38:** 1779-1788.

Mcdonald MB JR (1997) New techniques for cultivar identification of soybeans. *Seed Technology,* pp.117-130, Kalyani Publishers, New Delhi.

Naik BS (1998) Genetic characterization of cultivars and seed protein in mungbean. Unpublished. Ph.D. Thesis, Utkal University, Bhubaneshwar.

Padmavathi G, C Kole and EA Siddiq (1999) Detection of protein markers for identification of rice genotypes resistant to green leaf hopper. *Indian J. Genet.* **59:** 417-421.

Roy M, N Mandal and PK Das (2001) Seed protein characterization and isozyme diversity for cultivar identification in grass pea (*Lathyrus sativus* L.). *Indian J. Genet.* **61:** 246-249.

Sampat DS (1981) Chemotaxonomy of proteins of vigna species. Thesis Abst. **7:** 308-309.

Sasek A, J Bradora and J Cernl (1995) Hordein marker gene in certified Czech and Slovak varieties of barley (*Hordeum vulgare* L.). *Potravinosske vedy* **13:** 57-69.

Singh M, AP Singh and V Singh (2002) Identification for morphologically distant genotypes of *Catharanthus rosesus* by SDS-PAGE and peroxidase of leaf proteins. *J. Medicin. Aromat. Pl. Sci.* **24:** 68-73.

Stegemann H and G Pietsch (1983) Methods for characterization of seed proteins in cereals and legumes. In : W. Gottschack and H.P. Mueller (eds.), Seed Proteins : Biochemistry, Genetics, Nutritive Value, pp.45-75, Martinus Nijhaff, The Netherlands.

Wagner CK and MB Mcdonald (1982) Rapid laboratory tests useful for differentiation of soybean (*Glycine max*) cultivars. *Seed Sci. & Technol.* **10:** 431-439.