ACTA UNIVERSITATIS SAPIENTIAE
AGRICULTURE AND ENVIRONMENT, **12** (2020) 58–69

DOI: 10.2478/ausae-2020-0005

Genetic variation in common bean (*Phaseolus vulgaris* L.) using seed protein markers

Henok AYELIGN,[1,3] Eleni SHIFERAW,[2*] Faris HAILU[3]

[1] Worabe University, P. O. Box 89, Silte, Ethiopia, e-mail: ayelignhenok2005@gmail.com

[2] Ethiopian Biodiversity Institute, P. O. Box 30726, Addis Ababa, Ethiopia,
e-mail: eleni.shiferaw@ebi.gov.et

[3] College of Natural Science, Wollo University, P. O. Box 1145, Dessie, Ethiopia,
e-mail: markhmets@yahoo.com

Manuscript received: 10 April 2020; revised: 01 May 2020; accepted: 30 May 2020

Abstract. The genetic diversity of common bean accessions were assessed using seed storage protein markers. At regional level, accessions from the two major growing regions showed the highest level of gene diversity (H = 0.322, I = 0.485, and H = 0.312, I = 0.473), which can be exploited for the future improvement of the crop. Based on phaseolin, the major storage protein in common bean, the majority of the accessions (86%) were grouped under Mesoamerican gene pool. Seed proteins were also used to differentiate various Phaseolus species, indicating the usefulness of seed storage proteins in species identification in this genus.

Keywords: diversity, phaseolin, SDS-PAGE, seed protein

1. Introduction

Common bean (*Phaseolus vulgaris* L.) is among the cultivated bean types belonging to the genus Phaseolus. Due to long storage life, good nutritional properties, and requirements of easy storage and preparation, common bean has high value in the developing world [16]. It contains proteins (15%), starch (80%), and fats (2%); it provides folic acid, dietary fibre, and complex carbohydrates; some parts of the plant also have a therapeutic value [23].

Common bean is one of the major export commodities and a cash crop for small-scale farmers in Ethiopia, and it supplies about 60% of the total export of pulses [25]. Produced by about 3.38 million smallholders, it covered 21.6% of the total pulse growing area in the 2015/16 cropping season. The regions of

Oromia, Southern Nations Nationalities and Peoples' Region (SNNP), and Amhara are the major growing areas in the country [7]. Pests and diseases, lack of access to improved germplasms, and unreliable climatic conditions are among the constraints encountered by smallholder farmers that result in the low agricultural productivity of common bean [4].

Studying the diversity in common bean is useful to generate information that can be used in crop improvement and genebank management. Genetic diversity in common bean has been carried out using morphological markers [6], isozymes, seed proteins [3, 24], and various types of molecular markers [1, 2, 10, 18].

Seed storage protein fractions are mixtures of components which show polymorphism both within and among genotypes of the same species [22]. Bean seeds contain 20% to 25% proteins, dominated by the storage protein phaseolin [4], which determines both the quantity and nutritional quality of proteins in bean seeds [5, 12]. Polymorphism, environmental stability, and biochemical complexity characteristics enable phaseolin to be an informative marker [11]. Genetic diversity of common bean germplasms from Ethiopia has been carried out using inter-simple sequence repeat (ISSR) markers [8] and simple sequence repeat (SSR) markers [10]. To the knowledge of the authors, seed storage protein and phaseolin types have not been utilized to study Ethiopian Phaseolus collections. Hence, the present study was undertaken to analyse the suitability of seed storage protein for diversity assessment in Phaseolus collections of Ethiopia. It also examines the potential of seed storage proteins in discriminating different Phaseolus species and identifies phaseolin types observed in the analysed accessions.

2. Materials and methods

A total of 50 common bean accessions obtained from the Ethiopian Biodiversity Institute were used for this study (*Table 1*). Three accessions (240523, 235507, and 235506) conserved as *P. acutifolius*, one accession each conserved as *P. lunatus* (211481) and *P. sativus* (241742) were also included in the analysis. Fifteen seeds per accession were ground to fine powder with mortar and pestle. Seed proteins were extracted using 0.002M borate buffer. Protein profiling of extracted samples was analysed using sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) in 10% polyacrylamide gel [13]. Electrophoresis was carried out at constant voltage (100 V). At the end of the run, gels were stained with staining solution (40% ethanol, 10% acetic acid) containing 0.1% (w/v) Coommassie brilliant blue R-250 for overnight. Destaining of gels was carried out using 40% ethanol and 10% acetic acid.

The presence (1) or absence (0) of every band was scored in reference to a standard protein ladder (NEB P7712S). The resulting binary data matrix for the 50 accessions (750 individuals) was used to perform the data analysis. GenAlEx

version 6.5 [19] was used to calculate the percentage of polymorphic bands, heterozygosity, and diversity index. The relationship among the analysed samples was revealed by cluster analyses from the binary data using NTSYS v 2.1 [21]. To complement the information on clustering, principal coordinate analysis (PCoA) was conducted using GenAlEx software [19].

3. Results and discussions

Genetic diversity

On the basis of the relative mobility of seed proteins on the gel, a total of 41 reproducible bands were detected. The size of the protein bands detected ranged from 11 to 210 kDa. The Phaseolus accessions under study showed variations in total number of bands, which ranged from 14 to 33. The percentage of polymorphic bands ranged from 2.44% (accession 5) to 58.54% (accession 8) and averaged 27.05%, which revealed a wide and diverse genetic base in common bean accessions collected from different regions. High percentages of polymorphic loci were observed in Ac8, which was collected from the Oromia region (Mirab Harerge zone) and Ac41 from the Benishangul-Gumuz region (Metekel zone) with a value of 58.54% and 51.22% respectively (*Table 1*). The least band polymorphism was found in Ac5 (2.44%) collected from SNNP (Sidama zone). The highest gene diversity was shown by Ac8 (H = 0.213, I = 0.318), and the least gene diversity was observed in Ac5 (H = 0.002, I = 0.004). In general, at accession level, the diversity estimates show low level of variation within each accession (mean I = 0.135, mean H = 0.090). Using SSR markers, previous studies reported higher level of diversity among accessions collected from different regions of Ethiopia [1, 10], showing the limited potential of seed protein markers in revealing variation within each accession.

Table 1. Summary of parameters for genetic diversity assessment in common bean accessions: Accession code (code); Accession number (Ac No.); Region code (RC); percentage of polymorphic bands (PB), Heterozygosity (H), and diversity index (I)

Code	Ac No.	RC	PB	H	I	Code	Ac No.	RC	PB	H	I
Ac1	241756	1	17.07	0.058	0.087	Ac26	211338	2	7.32	0.020	0.031
Ac2	241742	1	26.83	0.103	0.150	Ac27	214664	1	4.88	0.014	0.020
Ac3	241737	1	17.07	0.063	0.094	Ac28	244805	1	7.32	0.033	0.047
Ac4	241738	1	41.46	0.175	0.253	Ac29	207935	3	24.39	0.084	0.126
Ac5	241733	1	2.44	0.002	0.004	Ac30	211277	1	26.83	0.064	0.103
Ac6	241734	1	17.07	0.049	0.075	Ac31	211282	1	21.95	0.057	0.089
Ac7	237080	2	9.76	0.048	0.067	Ac32	211276	1	24.39	0.077	0.118
Ac8	241134	2	58.54	0.213	0.318	Ac33	211278	1	12.20	0.049	0.072
Ac9	237993	1	17.07	0.073	0.105	Ac34	211376	2	24.39	0.085	0.128
Ac10	241736	1	41.46	0.107	0.167	Ac35	211266	5	31.71	0.090	0.140
Ac11	241748	1	17.07	0.048	0.075	Ac36	208639	2	31.71	0.091	0.142
Ac12	212861	2	26.83	0.089	0.134	Ac37	211322	2	26.83	0.092	0.138
Ac13	241739	1	31.71	0.102	0.155	Ac38	208637	2	43.90	0.138	0.210
Ac14	237079	2	24.39	0.084	0.126	Ac39	212860	2	19.51	0.059	0.090
Ac15	230044	2	36.59	0.147	0.214	Ac40	211313	2	26.83	0.084	0.128
Ac16	215048	2	31.71	0.101	0.153	Ac41	211344	3	51.22	0.150	0.230
Ac17	207943	2	41.46	0.140	0.211	Ac42	208645	2	17.07	0.053	0.082
Ac18	207940	3	36.59	0.086	0.138	Ac43	208705	2	39.02	0.114	0.174
Ac19	222872	4	48.78	0.142	0.219	Ac44	211340	2	26.83	0.091	0.137
Ac20	228911	2	34.15	0.105	0.159	Ac45	207933	3	26.83	0.091	0.135
Ac21	214665	1	21.95	0.082	0.119	Ac46	228912	2	31.71	0.104	0.156
Ac22	208638	2	29.27	0.109	0.160	Ac47	201293	2	24.39	0.094	0.137
Ac23	211386	5	36.59	0.125	0.184	Ac48	211345	3	21.95	0.071	0.107
Ac24	211378	2	39.02	0.136	0.204	Ac49	211265	5	17.07	0.067	0.098
Ac25	211355	3	36.59	0.140	0.205	Ac50	235697	2	21.95	0.088	0.128
								Mean	**27.05**	**0.090**	**0.135**

Note: 1 – SNNP, 2 – Oromia, 3 – Benishangul – Gumuz, 4 – Gambella, 5 – Amhara

Diversity estimates obtained by grouping populations based on their geographic origin (region) revealed that accessions from Oromia and SNNP regions showed the highest percentage of polymorphic bands with a value of 97.56% each, followed by Benishangul-Gumuz (87.80%), Amhara (73.17%), and

the least one was from Gambela region (48.78%). The highest diversity estimates were shown in Oromia (H = 0.322, I = 0.485) and SNNP (H = 0.312, I = 0.473) regions. Accessions from Oromia and SNNP were highly diverse in all the variability measures.

This agrees with the findings of Fisseha et al. [10], who reported high diversity on common bean accessions collected from these regions using SSR markers.

Similarly high diversity on common bean accessions collected from these regions was reported using SSR markers [10]. Hence, these two regions – which are also the most important regions in terms of common bean production in Ethiopia [25] – could be the most important regions for identifying important genes for the breeding and improvement of common bean.

On the basis of pair-wise genetic similarity matrix, the most distantly related accessions were Ac20 and Ac47, which were collected from the Oromia region, Illubabbor and Mirab Harerge zones, respectively, with a value of 0.751, indicating a wide range among the analysed accessions (data not shown). The most closely related accessions were Ac14 and Ac29 (GD = 0.00), which were collected from the Oromia and Benishangul-Gumuz regions respectively. The analysis of genetic distance among accessions pooled by region of collection revealed that the most distantly related accessions were the ones from SNNP and Gambela with a value of 0.085, and the least one was between Oromia and SNNP with a value of 0.012. Genetic distance is an important parameter for germplasm improvement, allowing the exploitation of distantly related populations, which may result in vigorous varieties that combine the traits of the distantly related parents.

Grouping of accessions based on similarity can be seen in *Figure 1*. The analysed accessions can be divided into three main clusters with accession 20 from the Oromia region, Illubabor zone shown as an outlier at coefficient of about 0.56. Accession 14 from Oromia and 29 from the Benishangul-Gumuz regions were the most similar ones (genetic similarity (GS) = 1.0), followed by accession 42 and 46 both from the Oromia region (GS = 0.86). Group I and II can be further divided into subclusters.

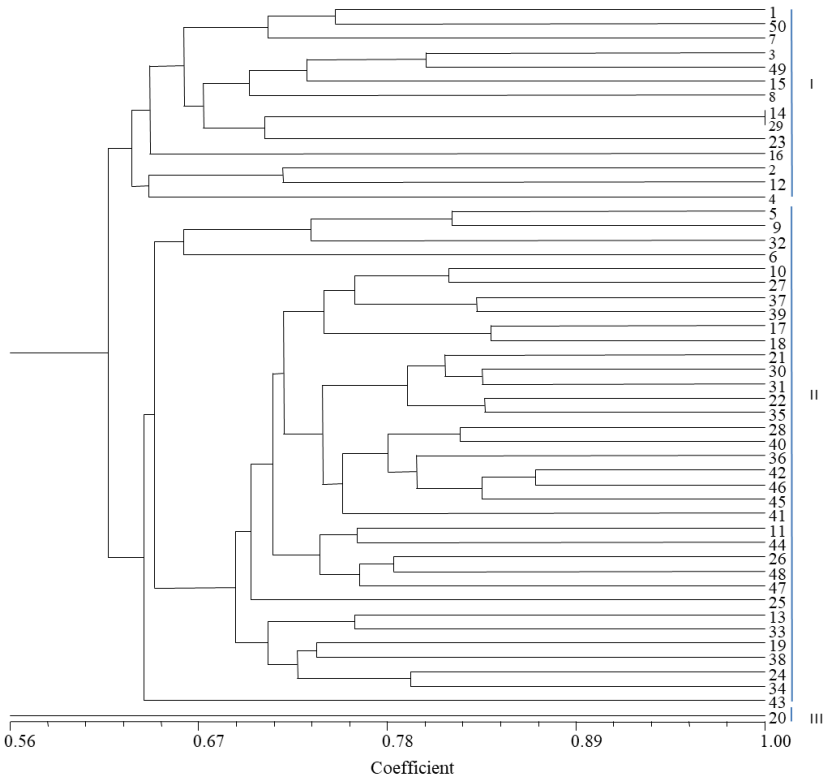

Figure 1. Dendrogram-compiled simple matching coefficient similarity coefficient showing the grouping of the analysed common bean accessions based on seed protein profile

The association among the analysed genotypes examined by principal coordinate analysis (PCoA) showed that the first three axes explained a cumulative variation of 60.03%. Individuals from an accession tend to group together; however, accessions from the same regions or zones did not show a specific grouping pattern (*Fig. 2*). In both PCoA and cluster analysis, accessions from different collection regions were grouped together, which may imply the exchange of plant materials and the possibility of gene dispersal by seeds in different common-bean-growing regions and zones of Ethiopia.

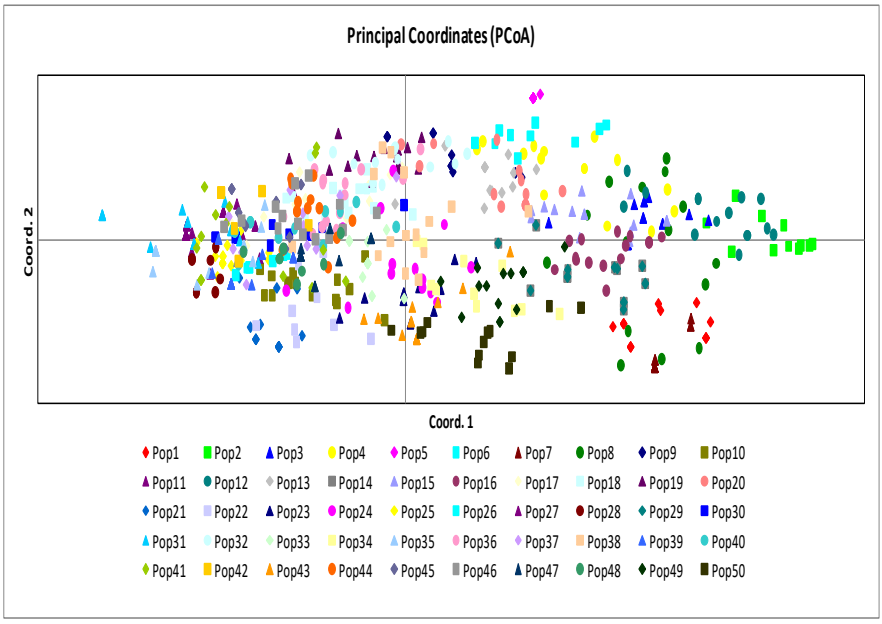

Figure 2. Principal coordinate analysis of 50 common bean accessions analysed by seed storage proteins

Diversity in phaseolin patterns

The protein profiles were also used to discriminate the existing variability by phaseolin patterns. Studies on phaseolin type determination allow researchers to understand the range of dispersal of bean genotypes from their original location to their secondary location [15]. Phaseolin protein has a narrow range of molecular weight (42–53 kDa), as reported by different researchers. The molecular weight for Phaseolin zone in the present study ranged from 42 kDa to 49 kDa (*Fig. 3*), which agrees with the findings reported by Madakbas et al. [15] and Tomlekova et al. [24]. This zone is used to determine the grouping of the accessions among the two centres of origins – Mesoamerican and Andean –, which are considered to be the primary centres of origin for common bean [11, 12].

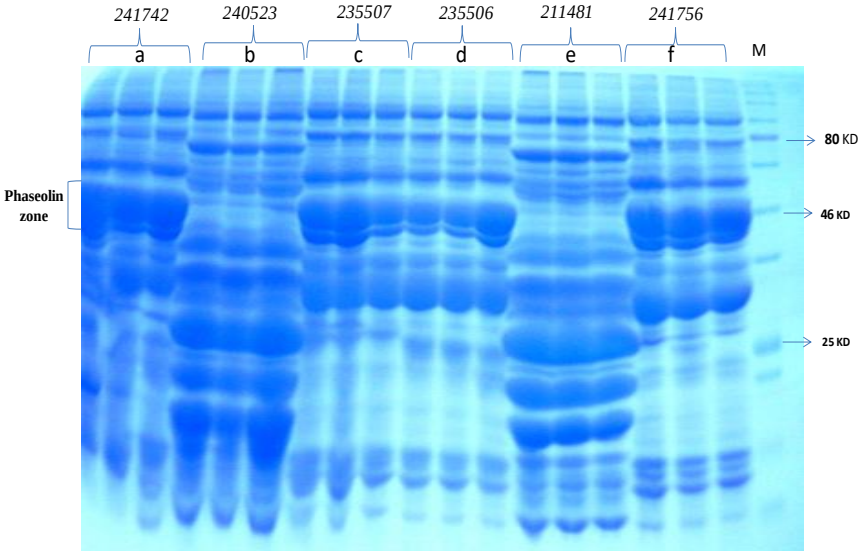

Figure 3. Similarity/difference among phaseolus species using seed protein profile: P. vulgaris (a, c, d), P. lunatus (b, e), f-control (P. vulgaris), M-molecular weight marker

There were three variants on this zone: three bands between 42 kDa and 49 kDa in the Andean phaseolin type (*Fig. 4*, samples c, d, e) and two bands in the Mesoamerican phaseolin type. The majority of the accessions analysed in this study were grouped under the Mesoamerican gene pool (86%) since they contained Mesoamerican phaseolin type, and seven accessions (Ac16, Ac17, Ac18, Ac32, Ac35, Ac36, and Ac44) contained samples that showed the Andean phaseolin type. According to Asfaw et al. [1], both the Mesoamerican and Andean gene pools are present in Ethiopia, with a higher frequency for the Mesoamerican type. Among East African countries, Andean genotypes are dominant in Kenya [1], while in Uganda the Andean and Mesoamerican genotypes occur with similar frequency [17]. The existence of both gene pools and the low proportion of Andean type common beans in Ethiopia may be due to the original introductions, subsequent imports of novel germplasm from various sources, the low level of adaptation of Andean types to the ecological conditions in Ethiopia, consumer preferences, and the occurrence of biotic and abiotic stresses which did not favour the Andean types [1, 27].

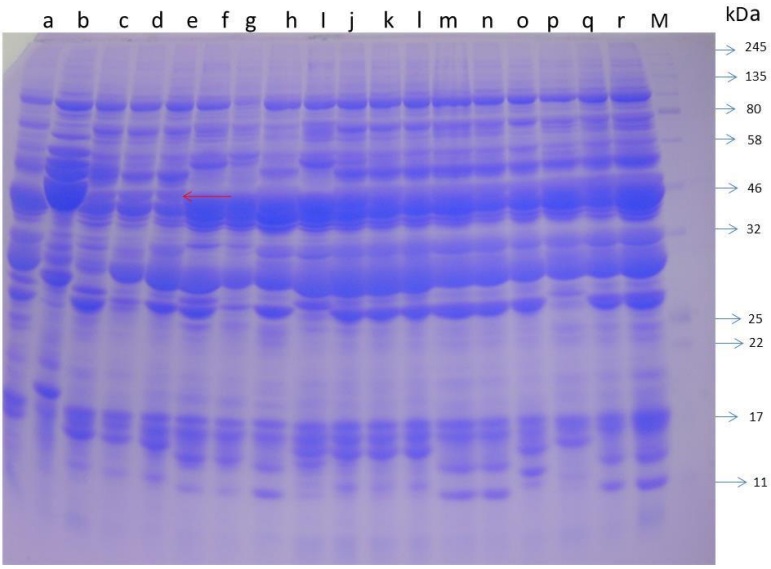

Figure 4. SDS-PAGE of total seed protein extracts of P. vulgaris samples. Samples c, d, and e show T-phaseolin patterns (indicated by arrow). M – protein weight marker

Assessment of the application of seed proteins for species identification in the phaseolus collection showed different banding patterns among the different species (*Fig. 3*). Three accessions were conserved as *P. sativus* and *P. acutifolius*. However, their banding pattern was similar to that of *P. vulgaris*. One accession conserved as *P. acutifolius* showed similar seed protein banding pattern with *P. lunatus*. These results were also supported by seed morphology and the examination of banding patterns using ISSR markers. Studies based on the electrophoretic analysis of seed proteins have been used to discriminate species and cultivars in other legumes as well. Seed protein profiles of different Lathyrus species showed unique electrophoresis patterns [9], and different electrophoretic seed albumin patterns were observed in different Lathyrus species [20]. In *Lupinus albus*, glutelins and glycoproteins were successfully used for cultivar identification [26], indicating the possible use of seed storage proteins in species identification in phaseolus and other legume species.

4. Conclusions

This study has examined genetic diversity in common bean using seed storage proteins. These markers show a low level of diversity within accessions, but a substantially higher level of diversity was observed among regions, which can be exploited for genetic improvement and further germplasm collection of the crop. Species identification using only plant morphology could result in misclassification of species. The study has also demonstrated that seed storage proteins, which are relatively inexpensive markers, can be employed in the identification of phaseolus species and genebank management.

Acknowledgments

This research was financially supported by the Ethiopian Biodiversity Institute.

References

[1] Asfaw, A., Blair, M. W., Almekinders, C. (2009), Genetic diversity and population structure of common bean (*Phaseolus vulgaris* L.) landraces from the East African highlands. *Theoret. Appl. Genet.* 120, 1–12.

[2] Becerra, V., Gepts, P. (1994), RFLP Diversity of Common Bean (*Phaseolus vulgaris*) in its centers of origin. *Genome* 37, 256–263.

[3] Berber, I., Yasar, F. (2011), Characterization of bean (*P. vulgaris* L.) cultivars grown in Turkey by SDS-PAGE of seed proteins. *Pak. J. of Bot.* 43(2), 1085–1090.

[4] Birachi, E. A., Ochieng, J., Wozemba, D., Ruraduma, C., Niyuhire, M. C., Ochieng, D. (2011), Factors influencing smallholder farmers' bean production and supply to market in Burundi. *Afr. Crop Sci. J.* 19(4), 335–342.

[5] Bliss, F. A., Brown J. W. S. (1983), Breeding common bean for improved quantity and quality of seed protein. In: Janick, J. (ed.), *Plant Breeding Reviews* 1. Boston, MA: Springer. 59–102.

[6] Chiorato, A. F., Carbonell, S. A. M., Dias, L. A., Moura, R. R., Chiavegato, M. B., Colombo, C. A. (2006), Identification of common bean (*Phaseolus vulgaris* L.) duplicates using agromorphological and molecular data. *J. Genet. Mol. Biol.* 29(1), 105–111.

[7] CSA. (2016), Agricultural sample survey 2015/2016. Volume I. Report on area and production of major crops (private peasant holdings, meher season). *Statistical Bulletin* 584. Addis Ababa, Ethiopia: Central Statistical Agency.

[8] Dagnew, K., Haileselassie, T., Feyissa, T. (2014), Genetic diversity study of common bean (*Phaseolus vulgaris* L.) germplasm from Ethiopia using inter simple sequence repeat (ISSR) markers. *Afri. J. Biotechnol.* 13(36), 3638–3649.

[9] Emre I., Turgut-Balik D., Sahin A. (2007), Electrophoretic analysis of total protein profiles of some Lathyrus L. (sect. Cicercula) grown in Turkey. *Pak. J. Biol. Sci.* 10, 2890–2894.

[10] Fisseha, Z., Tesfaye, K., Dagne, K., Blair, M. W., Harvey, J., Kyallo, M., Gepts, P. (2010), Genetic diversity and population structure of common bean (*Phaseolus vulgaris* L.)

germplasm of Ethiopia as revealed by microsatellite markers. *Afr. J. Biotechnol.* 15(52), 2824–2847.

[11] Gepts, P. (1988), Phaseolin as an evolutionary marker. In: Gepts, P. (ed.), Genetic resources of Phaseolus beans. *Current Plant Science and Biotechnology in Agriculture* 6, 215–241. Dordrecht: Springer.

[12] Gepts, P., Osborn, T. C., Rashka, K., Bliss, F. A. (1986), Phaseolin protein variability in wild forms and landraces of common bean (*Phaseolus vulgaris*): Evidence for multiple centers of domestication. *Econ. Bot.* 40, 451–486.

[13] Lammeli, U. K. (1970), Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680–685.

[14] Ma, Y., Bliss, A. F. (1978), Seed proteins of common bean. *Crop Sci.* 18, 431–437.

[15] Madakbas, S. Y., Awale, H., Kelly, J. (2014), Determination of Phaseolin types in common bean (*Phaseolus vulgaris*) varieties from Turkey. *Greener J. Agric. Sci.* 4(2), 39–45.

[16] Namugwanya, M., Tenywa, J. S., Otabbong, E., Mubiru, D., Basamba, T. (2014), Development of common bean (*Phaseolus vulgaris* L.) production under low soil phosphorus and drought in sub-Saharan Africa: A review. *J. Sustain. Dev.* 7(5), 128–139.

[17] Okii, D., Tukamuhabwa, P., Kami, J., Namayanja, A., Paparu, P., Ugen, M., Gepts, P. (2014), The genetic diversity and population structure of common bean (*Phaseolus vulgaris* L) germplasm in Uganda. *Afri. J. Biotechnol*. 13 (29), 2395–2949.

[18] Palomino, E. C., Mori, E. S., Zimback, L., Tamborussi, E. V., Moraes, C. B. (2005), Genetic diversity of common bean genotypes of Carioca commercial group using RAPD markers. *Crop Breed. Appl. Biotechnol.* 5, 80–85.

[19] Peakall, R., Smouse, P. E. (2006), GenAlEx 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Mol. Ecol. Notes* 6, 288–295.

[20] Przybylska, J., Zimniak Przybylska Z., Krajewski, P. (1999), Diversity of seed albumins in some Lathyrus species related to *L. sativus* L.: An electrophoretic study. *Genet. Res. Crop Evo.* 46, 261–266.

[21] Rohlf, F. J. (2000), NTSYSpc, *Numerical Taxonomy and Multivariate Analysis System, Version 2.1*. New York, USA: Exeter Publications.

[22] Shewry, P. R., Napier, J. A., Tatham, A. S. (1995), Seed storage proteins: Structures and biosynthesis. *Plant Cell* 7, 945–956.

[23] Svetleva, D., Pereira, G., Carlier, J., Cabrita, L., Leitao, J., Genchev, D. (2006), Molecular characterization of *Phaseolus vulgaris* L. genotypes included in Bulgarian collection by ISSR and AFLP analyses. *Sci. Hort.* 109, 198–206.

[24] Tomlekova, N., Sofkova, S., Sarsu, F., Baudoin, J. P. (2016), Genetic diversity of Bulgarian *Phaseolus vulgaris* L. based on Phaseolin type and seed coat colour. *Bulg. J. Agric. Sci.* 22, 447–451.

[25] Van den Broek, J., Sertse, Y., Becx, G., Paulos, A., Tesfaye, B., Zenaw, D., Teshome, K. (2014), *Legume value chains in Ethiopia: Landscaping study*. Available at: https://www.agriknowledge.org/file_downloads/g158bh34t (accessed on: 12 April 2018).

[26] Vaz, A. C., Pinheiro, C., Martins, J. M. N., Ricardo, C. P. P. (2004), Cultivar discrimination of Portuguese *Lupinus albus* by seed protein electrophoresis: The importance of considering "glutelins" and glycoproteins. *Field Crop Res.* 87, 23–34.

[27] Wortmann, S. C., Kirkby, A. R., Eledu, A. C., Allen, J. D. (1998), Atlas of common bean (*Phaseolus vulgaris* L.) production in Africa. *Pan-African Bean Research Alliance* 133. Cali, Colombia: International Centre for Tropical Agriculture, CIAT.