



Genetics and Breeding

Statistical analysis of genetic diversity using faba bean landraces database

N. Velcheva*, S. Petrova

Agricultural Academy, Institute of Plant Genetic Resources K. Malkov, 2 Druzhiba Str., 4122 Sadovo, Bulgaria

(Manuscript received 20 May 2020; accepted for publication 24 July 2020)

Abstract. Evaluation of genetic diversity among landraces could be an invaluable aid related to the sustainable use of *ex situ* collections. Statistical methods are currently available for analysis of databases from investigation of stored germplasm. Faba bean (*Vicia faba* L.) is a self-pollinating with high percentage of foreign pollination legume crop with a great importance for food and forage due to its high protein content as well as the important role in soil fertility and nitrogen fixation. The local populations are well adapted to specific agro-climatic conditions in the growing areas and are a rich initial material for the breeding programs. The purpose of this study is to establish the genetic diversity of 21 Bulgarian faba bean landraces by important traits in order to review the current potential of conserved germplasm for its sustainable use. All genotypes, included in the study, are collected from expeditions in the country, recorded in the National Register for Plant Genetic Resources and long term stored at the National Genebank. They are characterized according to the International Faba Bean Descriptors. The cluster analysis results show a high genetic diversity in the collection and the variability of each studied trait is presented. The factor analysis, which complements the cluster analysis, gives a reason to group the genotypes with their features into groups that have a breeding value. Genetic diversity in the studied collection has been identified and some of the landraces could be included in future breeding programs.

Key words: *Vicia faba* L., passport data, characterization, evaluation, long term storage

Introduction

Traditional varieties, generally known as landraces, have been continuously maintained by farmers within their biological, cultural and socio-economic context. The local populations are well adapted to specific agro-climatic conditions in the growing areas and are a rich initial material for the breeding programs (Krasteva et al., 2009; Negri et al., 2009). Natural agro-biodiversity stored in genebanks can be used to enhance the diversity of cultivated plants with trait variation that confer tolerance to biotic and abiotic stresses due to the climate changes (Maxted et al., 2000; Khazaei et al., 2013).

Faba bean is a self-pollinating with high percentage of foreign pollination legume crop with a great importance for food and forage due to its high protein content as well as the important role in soil fertility and nitrogen fixation. The *Vicia faba* L. collections are represented only by cultivated forms, which make crucial preservation of local varieties for the present and future breeding activities of this crop, especially due to the absence of a natural reservoir of wild accessions and of the modernization of agriculture, which progressively phases out numerous landraces. Large investments in the discovery of genetic variability and also in breeding activity for traits of agronomic interest have been

made for faba bean at the end of the 20th century in European countries and also at ICARDA Syria (Duc, 1997; Ambrose et al., 2006; Link, 2006). Farmer's populations are particularly appropriate for the proper maintenance of genetic variation that is initially present in accessions (Duc et al., 2010; Alghamdi et al., 2012; Yilmaz, 2020).

Evaluation of genetic diversity among landraces could be an invaluable aid related to the sustainable use of *ex situ* collections. Statistical methods are currently available for analysis of databases from investigation of stored germplasm. Multivariate methods such as cluster and factor analysis can be performed on morphological quantitative data as well as the mean, range and standard deviation for determining of variation of the traits could be discussed and explaining their effects in the final output of clustering (Mohammadi and Prasanna, 2003). For proper selection of genotypes, involved in the breeding programs and for reduction of the improvement process, the breeder needs a preliminary study of the genetic diversity in the stored germplasm collections. It is necessary to use genetically distant genotypes for successfully combining the important features in the new genome (Stamatov, 2015).

The purpose of this study is to establish the genetic diversity of Bulgarian faba bean landraces by important traits in order

*e-mail: nikolaya_velcheva@abv.bg

to review the current potential of conserved germplasm in the National Genebank for its use in future breeding programs.

Material and methods

Plant material and studied traits

During the period 2018-2019 at the Institute of Plant Genetic Resources – Sadovo experimental field, 21 *Vicia faba* L. landraces were studied according to the Faba Bean Descriptors of IBPGRI (1985) and UPOV (2002).

The genotypes are collected by expeditions from small rural and suburban farms in the different regions in the country (Table 1). They are recorded with full agro ecological passport description in the National Register for Plant Genetic Resources. The information includes: FAO Institute code (Instcode), year of collection, catalogue number, taxonomic description, accession name (Accename), complete ecological and geographical characteristics of the location of origin and collecting source, biological status, as well as the type of germplasm storage.

Table 1. Passport data of *Vicia faba* local accessions, included in the study

No	Instcode	Year	Cat. N	Collecting site	Latitude	Longitude	Elevation
1	BGR001	1986	86E2179	v. Gramatikovo, Burgas	423043 N	273917 E	295 m
2	BGR001	1986	86E2180	t. Lubimets, Haskovo	415000 N	265000 E	56 m
3	BGR001	1986	86E2182	v. Pastrogor, Svilengrad	415100 N	261200 E	120 m
4	BGR001	1986	86E2183	v. Raikova mogila, Haskovo	414900 N	261800 E	188 m
5	BGR001	1986	86E2184	v. Mezek, Haskovo	414400 N	265000 E	167 m
6	BGR001	1986	86E2185	v. Raikova mogila, Haskovo	414900 N	261800 E	188 m
7	BGR001	1986	86E2186	v. Levka, Haskovo	415215 N	261600 E	217 m
8	BGR001	2008	A8E0001	v. Stezherovo, Pleven	433000 N	259000 E	130 m
9	BGR001	2008	A8E0049	v. Kochevo, Plovdiv	426010 N	245512 E	166 m
10	BGR001	2008	A8E0050	v. Nevsha, Varna	431600 N	271800 E	114 m
11	BGR001	2009	A9E0289	v. Hitovo, Dobrich	434700 N	273400 E	222 m
12	BGR001	2009	A9E0301	v. Hitovo, Dobrich	434700 N	273400 E	222 m
13	BGR001	2009	A9E0521	v. Tsenovo, Russe	433200 N	253900 E	50 m
14	BGR001	2010	B0E0001	v. Sinitovo, Pazardzhik	429000 N	242200 E	199 m
15	BGR001	2010	B0E0002	v. Sinitovo, Pazardzhik	429000 N	242200 E	199 m
16	BGR001	2010	B0E0003	v. Sinitovo, Pazardzhik	429000 N	242200 E	199 m
17	BGR001	2010	B0E0005	v. Karadzhevo, Plovdiv	426000 N	245400 E	165 m
18	BGR001	2010	B0E0006	t. Suvorovo, Varna	431948 N	273533 E	246 m
19	BGR001	2010	B0E0011	t. Yambol	422940 N	263031 E	114 m
20	BGR001	2010	B0E0012	v. Kozarevo, Yambol	422700 N	263600 E	1 m
21	BGR001	2010	B0E0016	t. Troyan, Lovech	425340 N	244234 E	380 m

*Data source- National Register of Plant Genetic Resources of IPGR Sadovo; v- village; t- town

All genotypes are long term stored at the National Genebank in hermetically closed bags at -18°C and at 3-7% moisture. The Bulgarian *Vicia faba* L. collection is published and available with open access in the electronic catalogue on Plant Genetic Resources EURISCO (<http://eurisco.ipk-gatersleben.de>).

Database of the results from the study of the quantitative traits: plant height (cm), branching (number), number of pods per plant, number of seeds per plant, number of seeds per pod, weight of seeds per plant (g), mass of 100 seeds (g), duration of flowering (days), vegetation period (days), was created.

Statistical data analysis

The obtained results were processed using statistical package SPSS 19.0. for Windows (IBM SPSS Statistics 19 Product Version: 19.0.0). Variability of the study traits through the values of the coefficient of variation and cluster analysis of the experimental data were made. It was accepted that

the variation is considered to be weak if the coefficient of variation does not exceed 10%, middle - when it is greater than 10% and less than 20%, strongly (significant) - when it is above 20% (Dimova and Marinkov, 1999). The genetic distance of the traits was determined using cluster analysis. The accessions are grouped according to the Euclidean distance. The cluster analysis is successfully supplemented by factor analysis by the method of Principal Component Analysis for use of studied materials in crop breeding (Dimova and Bojinov, 2001). It determines the significant components that unite the studied traits and on this basis a breeding program for increasing their variability could be done.

Results and discussion

One of the main reasons that determines the high diversity of Bulgarian landraces is that the country contains

three phyto-climatic areas: Central-European, Mediterranean and Euro-Asian steppe and forest regions. The analysis of the passport database shows that expeditions for local plant genetic resources, included in the study, were carried out in the regions of North and South Bulgaria, flat and semi-mountainous regions with an altitude of 1 to 380 m (Table 1). In home gardens still traditional varieties suitable for the conditions of the region in order to meet family needs and for local markets are produced.

The results of the statistical analysis show that certain studied traits are characterized by high variability and fall into different cluster groups. The calculated indicators

characterize the average value of the quantitative traits, as well as the variation coefficients (CV) (Table 2). All studied traits were characterized by variability above 10%. Middle variability (10-20%) is shown by plant height, number of seeds per pod, mass of 100 seeds and duration of flowering. High genetic diversity in the collection is influenced by the significant variability traits: number of branches (CV= 32.51%), number of pods per plant (CV= 33.45%), number of seeds per plant (CV= 25.54%), weight of seeds per plant (CV= 33.03%), and vegetation period (CV= 40.71%). These five traits are most responsible for the high genetic diversity in the collection.

Table 2. Experimental database according to the Faba Bean Descriptors

No	Cat. No	Plant height (cm)	Branching (number)	Number of pods per plant	Number of seeds per plant	Number of seeds per pod	Weight of seeds per plant (g)	Mass of 100 seeds (g)	Duration of flowering (days)	Vegetation period (days)
1	86E2179	74.300	2.400	9.800	23.000	2.200	33.550	153.100	28.000	87.000
2	86E2180	71.700	2.900	10.300	21.500	2.100	30.210	150.200	28.000	87.000
3	86E2182	68.000	2.900	12.300	27.100	2.200	30.910	132.400	33.000	89.000
4	86E2183	71.300	1.900	7.300	17.400	2.400	26.330	149.100	33.000	89.000
5	86E2184	69.000	2.700	12.400	23.400	1.900	28.160	127.500	33.000	89.000
6	86E2185	71.500	2.200	8.000	19.400	2.400	31.030	150.400	28.000	87.000
7	86E2186	67.900	2.700	14.100	30.400	2.100	34.520	152.100	33.000	89.000
8	A8E0001	76.500	2.800	7.100	16.900	2.400	22.850	143.720	42.000	198.000
9	A8E0049	92.100	2.600	9.000	30.200	3.400	42.750	148.760	37.000	198.000
10	A8E0050	73.900	2.100	5.400	21.300	3.900	38.700	177.240	32.000	198.000
11	A9E0289	96.200	2.900	16.700	34.600	2.100	60.550	149.500	41.000	102.000
12	A9E0301	93.000	2.000	13.200	32.700	2.500	32.710	116.500	41.000	91.000
13	A9E0521	72.600	3.300	7.100	17.000	2.400	22.730	129.500	41.000	91.000
14	B0E0001	75.700	5.900	14.200	34.200	2.500	58.340	156.900	31.000	91.000
15	B0E0002	83.300	3.700	11.700	30.400	2.600	49.170	150.320	28.000	91.000
16	B0E0003	74.500	4.900	16.000	32.100	2.000	59.160	182.600	28.000	91.000
17	B0E0005	73.500	3.100	6.300	18.000	2.900	31.540	191.160	32.000	198.000
18	B0E0006	97.800	1.900	9.700	29.700	3.100	23.310	77.800	35.000	201.000
19	B0E0011	68.500	2.900	6.200	20.700	3.000	33.810	202.100	26.000	91.000
20	B0E0012	77.800	3.600	7.800	21.100	2.800	37.970	172.200	31.000	91.000
21	B0E0016	71.500	2.900	7.600	17.000	2.200	24.160	150.720	31.000	91.000
Mean		77.171	2.967	10.105	24.671	2.529	35.831	150.658	32.952	116.190
Minimum		67.900	1.900	5.400	16.900	1.900	22.730	77.800	26.000	87.000
Maximum		97.800	5.900	16.700	34.600	3.900	60.550	202.100	42.000	201.000
Range		29.900	4.000	11.300	17.700	2.000	37.820	124.300	16.000	114.000
Std. Deviation		9.494	0.965	3.380	6.301	0.501	11.834	26.808	4.904	47.305
CV, %		12.303	32.513	33.451	25.539	19.819	33.026	17.794	14.882	40.713

From our previous studies we established that the following traits: plant height, number of pods per plant and weight of seeds per plant can be used as selection criteria to increase number of seeds per plant in *Vicia faba* genotypes (Petrova, 2019). The variability of these

four traits is represented by histograms visualizing the frequency of appearance on the basis of Standard Deviation (Figure 1). The graphs give an indication of presence of genetic diversity in the studied quantitative characteristics.

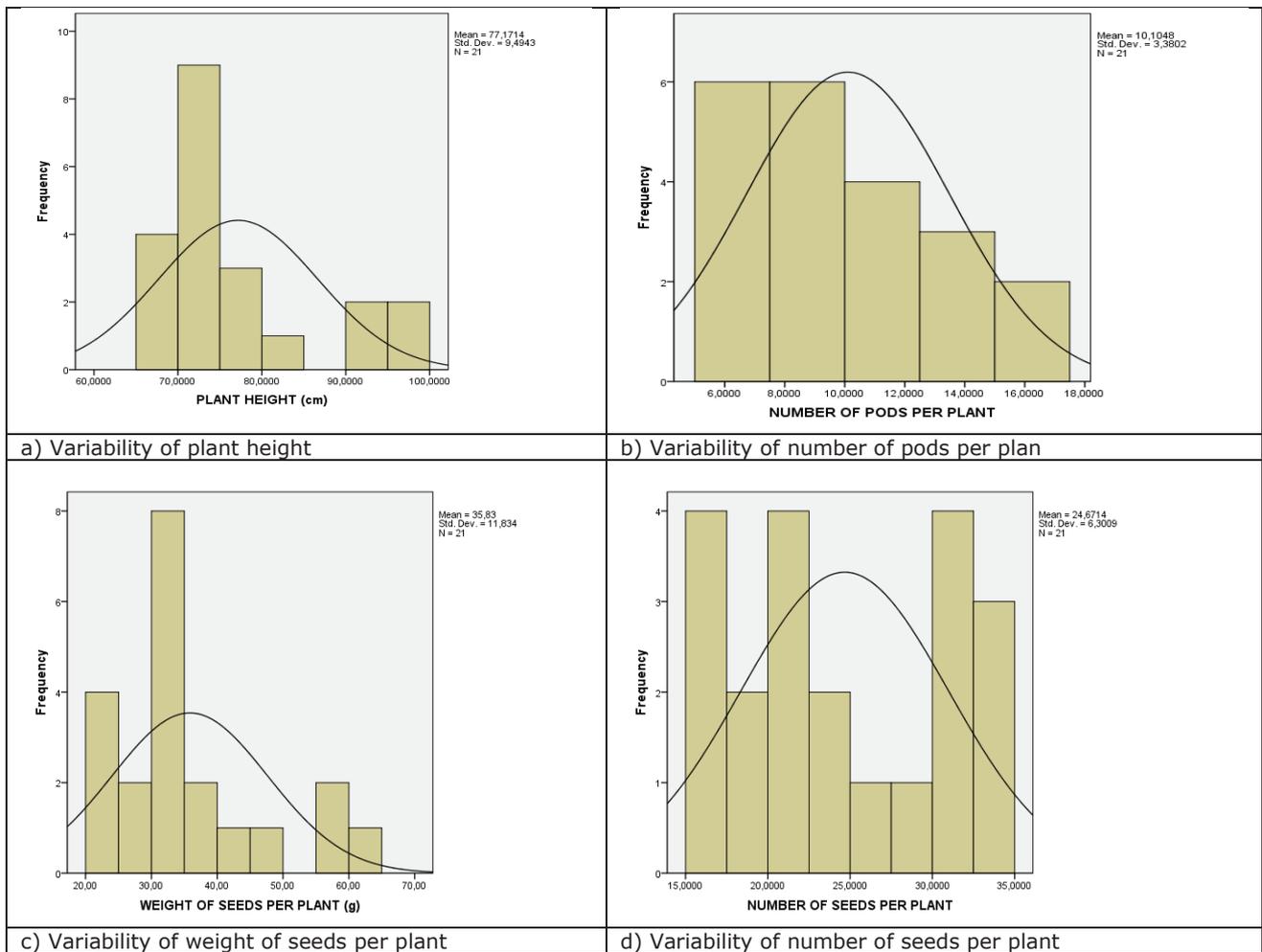


Figure 1. Histograms of the variability of some studied traits: a) Variability of plant height; b) Variability of number of pods per plant; c) Variability of weight of seeds per plant; d) Variability of number of seeds per plant

The conducted cluster analysis groups the genotypes into two clusters (Figure 2). The first cluster includes 16 genotypes with diverse geographical origin and short vegetation period (from 87 to 102 days). This group includes accession B0E0002, originating from Sinitovo, Pazardzhik region, which exceeds the average by eight of the studied traits. The accession A9E0289 (Hitovo, Dobrich region) is characterized by the highest values of the traits: number of pods per plant, number of seeds per plant and weight of seeds per plant in the collection, with great economic importance for the crop. The genotypes with cat. No. B0E0003 (Sinitovo, Pazardzhik region), B0E0012 and B0E0011 from Yambol region are characterized by a high number of seeds per pod and an equal vegetation period (91 days).

The second cluster includes five genotypes with diverse geographical origin and vegetation period from 198 to 201 days. The accession B0E0006, originating from Suvorovo, Varna region has the highest plant height (97.8 cm) and the longest vegetation period in the collection (201 days). This genotype falls into a separate forket.

From the clustering of the traits we can conclude that they are conditionally combined into two groups responsible for the grouping of the genotypes in the collection (Figure 3). The first cluster sequentially includes traits: number of

branches, number of seeds per pod, number of pods per plant, number of seeds per plant, duration of flowering and weight of seeds per plant, combined at a close Euclidean distance. The plant height, vegetation period and mass of 100 seeds are combined into a second cluster at a large Euclidean distance.

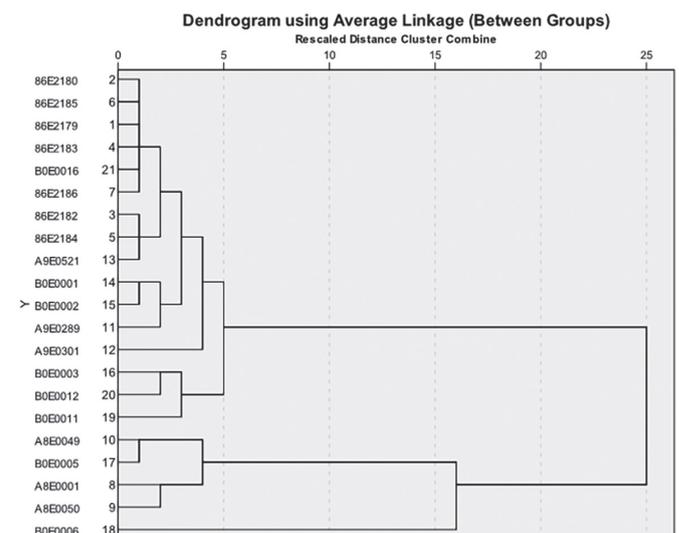


Figure 2. Dendrogram of cluster analysis of genotypes

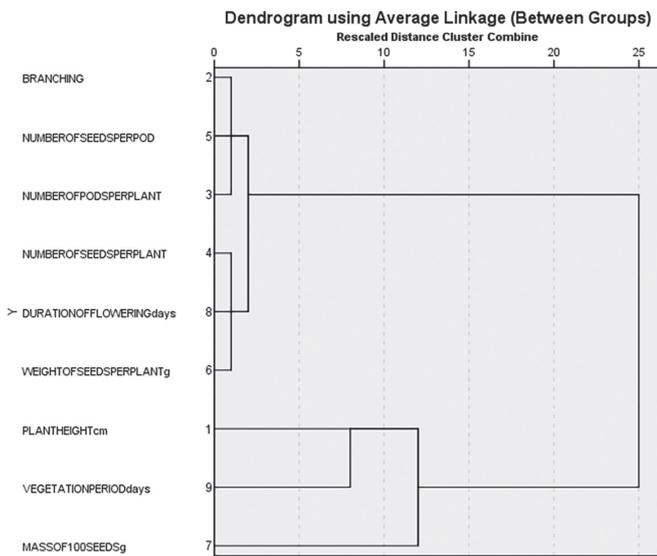


Figure 3. Dendrogram of cluster analysis of studied traits

Similar results with proven high variability in traits in local varieties from Turkey have been reported by Karaköy et al. (2014). As revealed by Terzopoulos et al. (2003), the grouping of *Vicia faba* genotypes, originating from Greece, is based mainly on pod characteristics, plant height, 1000 seed weight and branching from basal nodes.

The Principal component analysis separates the studied traits into three significant factors (Figure 4).

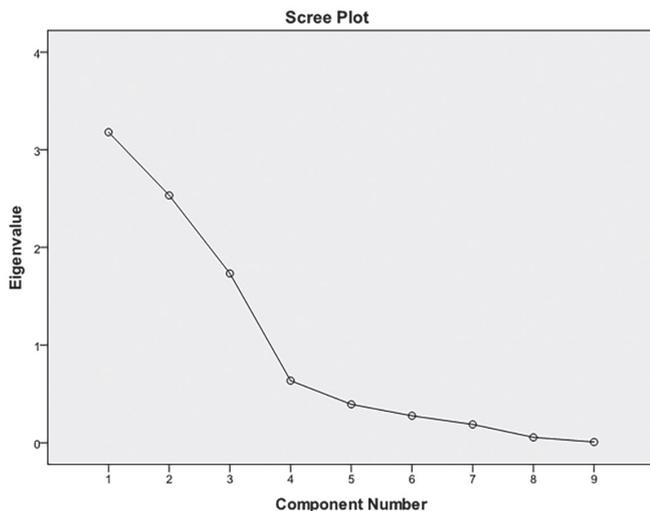


Figure 4. Factorial plane of significant components

The first significant factor includes number of branches, number of pods and seeds per plant and weight of seeds per plant. The second factor includes plant height, mass of 100 seeds, the flowering and vegetation period. The third significant factor separates the number of seeds per pod (Table 3, Figure 5). The collection in one factor of individual traits leads to their joint transmission in the offspring (Stamatov and Deshev, 2012). This is a fact that would also facilitate the breeding process in achieving the desired model in hybridization.

Table 3. Component matrix of Principal Component Analysis

Traits	Component		
	1	2	3
Plant height	-	0.874	-
Branching	0.650	-	-
Pods per plant	0.933	-	-
Seeds per plant	0.837	-	-
Seeds per pod	-	-	0.757
Weight of seeds per plant	0.825	-	-
Mass of 100 seeds	-	-0.683	-
Duration of flowering	-	0.739	-
Vegetation period	-	0.548	-

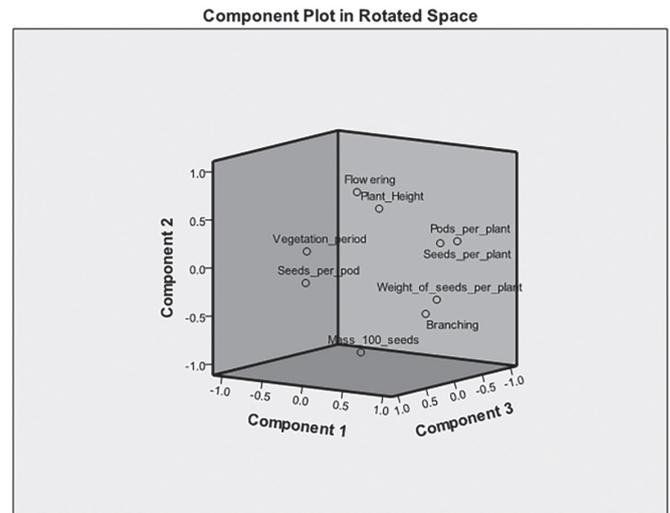


Figure 5. Component plot of Principal Component Analysis

The results from the conducted factor analysis (Figure 6) show that the most distant from the center of the factorial plane are accessions 11 (A9E0289), 14 (B0E0001), 15 (B0E0002) and 16 (B0E0003). They could be used as parents if they form parental pairs with 10 (A8E0050), 17 (B0E0005) and 19 (B0E0011). Using this result, high variability in the studied traits would be achieved in future hybrids.

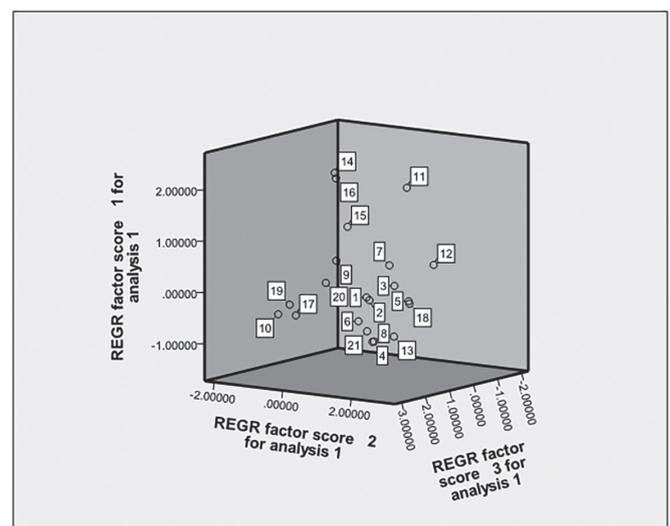


Figure 6. Results from factor analysis

Conclusion

Ex situ collection of 21 faba bean genotypes is characterized by nine quantitative traits. The variability of the studied traits has shown high genetic diversity in the studied germplasm collection. The database from the investigation complements the passport characterization of the *Vicia faba* L. accessions and increases the possibilities for sustainable use of the conserved germplasm collection in the National Genebank. The analysis grouped the collection in two clusters by the quantitative traits with economic importance for the crop. The traits are combined into two cluster groups, that would facilitate the crop breeding programs using genetically spaced genotypes as parental pairs. The results from the factor analysis give us the opportunity to determine the parental pairs in order to realize the desired model of a plant and obtain high variability in the studied traits. The proved diversity in Bulgarian faba bean landrace collection with different geographical origin makes the genotypes interesting genetic resource for future research activities.

Acknowledgements

This work was supported by the Bulgarian Ministry of Education and Science under the National Research Programme “Healthy Foods for a Strong Bio-Economy and Quality of Life”, approved by DCM No. 577/17.08.2018.

References

- Alghamdi SS, Migdadi HM, Ammar MH, Jeffrey GP and Siddique KHM**, 2012. Faba bean genomics: current status and future prospects. *Euphytica*, 186, 609-624.
- Ambrose M, Duc G, Angelova S and Suso MJ**, 2006. Report of the ECP/GR Oil and Protein Crops Network, In: Report of the ECP/GR Network coordinating groups: First meeting (eds. L. Maggioni and E. Lipman), Rome, Italy, pp. 54-73.
- Dimova D and Bojinov B**, 2001. Application of cluster analysis and principal component analysis for evaluation of selection materials. Anniversary scientific session “Selection and agrotechnology of field cultures”, Dobrich, I, 308-312.
- Dimova D and Marinkov E**, 1999. Experiment and Biometrics. Academic Publishing House of Agricultural University, Plovdiv, p. 194 (Bg).
- Duc G**, 1997. Faba bean (*Vicia faba* L.). *Field Crops Research*, 53, 99-109.
- Duc G, Bao S, Baum M, Redden B, Sadiki M, Suso MJ, Vishniakova M and Zong X**, 2010. Diversity maintenance and use of *Vicia faba* L. genetic resources. *Field Crops Research*, 115, 3, 270-278.
- IBM SPSS Statistics 19 Product Version: 19.0.0 Serial Number: NA.**
- IBPGRI**, 1985. *Faba Bean Descriptors*, Rome, Italy.
- Karaköy T, Baloch F, Toklu F and Özkan H**, 2014. Variation for selected morphological and quality-related traits among 178 faba bean landraces collected from Turkey. *Plant Genetic Resources*, 12, 5-13.
- Khazaei H, Street K, Santanen A, Bari A and Stoddard FL**, 2013. Do faba bean (*Vicia faba* L.) accessions from environments with contrasting seasonal moisture availabilities differ in stomatal characteristics and related traits? *Genetic Resources and Crop Evolution*, 60, 2343-2357.
- Krasteva L, Stoilova T, Varbanova K and Neykov S**, 2009. Bulgarian Landrace Inventory – Significance and use. *Biodiversity Technic. Bulletin, Biodiversity International*, 15, Rome, Italy, pp. 53-67.
- Link W**, 2006. Methods and objectives in faba bean breeding. In: *Proceedings of the International Workshop on Faba Bean Breeding and Agronomy* (eds. C.M. Avila, J.I. Cubero, M.T. Moreno, M.J. Suso and A.M. Torres), Spain, 35-40.
- Maxted N, Erskine W, Singh DP, Robertson LD and Asthana AN**, 2000. Are our germplasm collections museum items? In: *Proceedings of the Third International Food Legumes Research Conference: Linking Research and Marketing Opportunities for Pulses in the 21st Century*, Adelaide, Australia, 22-26 September 1997 (ed. R. Knight), Kluwer Academic Publishers, Dordrecht, The Netherlands.
- Mohammadi SA and Prasanna BM**, 2003. Analysis of genetic diversity in crop plants - Salient statistical tools and considerations. *Crop Science*, 43, 1235-1248.
- Negri V, Maxted N and Vetelainen M**, 2009. European Landrace Conservation: an Introduction. *Biodiversity Technical Bulletin, Biodiversity International*, Rome, Italy, 15, 1-22.
- Petrova S**, 2019. Selection criteria for increasing grain yield per plant of *Vicia faba* L. in Sadovo region, Bulgaria. *Journal of Balkan Ecology*, 22, 398-401.
- Stamatov S**, 2015. Peanuts selection program, choice of parents. *Agrarni Nauki (Agrarian sciences)*, Agricultural University – Plovdiv, 7, 43-49 (Bg).
- Stamatov S and Deshev M**, 2012. Evaluation of the possibilities for increasing the number of capsules in sesame (*Sesamum indicum* L.) by using the Factor analysis. *Plant Sciences*, 6, 81-84.
- Terzopoulos P, Kaltsikes P and Bebeli P**, 2003. Collection, evaluation and classification of Greek populations of faba bean (*Vicia faba* L.). *Genetic Resources and Crop Evolution*, 50, 373-381.
- UPOV**, 2002. Guidelines for the conduct of tests for distinctness, uniformity and stability of field bean (*Vicia faba* L. var. *minor*). International Union for Protection of New Varieties of Plants.
- Yilmaz N**, 2020. Morphological and molecular characterization of local faba bean (*Vicia faba* L.) accessions using inter-simple sequence repeat (ISSR) markers. *Fresenius Environmental Bulletin*, 29, 3756-3763.