# GENETIC DIVERSITY OF THE COMMON BEAN (*PHASEOLUS VULGARIS* L.) LANDRACES FROM BOSNIA AND HERZEGOVINA ASSESSED USING MICROSATELLITE MARKERS

Meliha Bajrić[1], Fuad Gaši[2], Arnela Okić[2], Mirza Hadžidedić[3], Almira Konjić[2], Jasmin Grahić[2]

Original scientific paper - *Originalni naučni rad*

**Summary**

This study aimed to analyze the genetic diversity of common bean (*Phaseolus vulagris* L.) landraces from the existing collection maintained at the Gene-bank of the Faculty of Agriculture and Food Sciences, University of Sarajevo using microsatellite markers (SSR). A total of 21 accessions of common bean were genotyped using ten SSR markers. The obtained molecular data was analyzed using genetic structure and factorial correspondence analyses. The examined accessions of common bean were not completely structured in accordance to the gene center of origin, which indicates that during the long cultivation period of common bean in Bosnia and Herzegovina (B&H), spontaneous crosses occurred between Mesoamerican and Andean common bean genotypes. The results of analyses confirmed that this was a heterogeneous and very interesting genetic material.

Key words: *common bean, SSR markers, gene-center of origin, analysis of genetic structure, factorial correspondence analysis*

## INTRODUCTION

The common bean (*Phaseolus vulgaris* L.) originated in Central America, more specifically – in todays Mexico, from where it spread to the south (Andes) (Bitocchi *et al.*, 2013). There are two major centers of origin of common bean, Mesoamerican and Andean (Bitocchi *et al.*, 2013; Cortés *et al.*, 2011; McClean *et al.*, 2004). In a study by Papa *et al.* (2006) the introduction and the path of expansion of common bean into Europe (France, Germany, Netherlands, and the Former Soviet Union) and centuries of bean cultivation in Bosnia and Herzegovina (B&H) have led to the development of traditional cultivars adapted to the specific climatic and edaphic conditions present in this area, as well as to the traditional cultivating methods.

B&Hs common bean landraces that were cultivated and thus conserved, by the efforts of small-scale farmers at various locations throughout the country, were collected during 2011 and 2012 and stored in the Gene-bank of the Faculty of Agriculture and

[1] Technical University of Munich, TUM School of Life Sciences Weihenstephan
[2] University of Sarajevo, Faculty of Agriculture and Food Sciences
[3] Poljoprivredno gazdinstvo Hadžidedić
Korespondencija: j.grahic@ppf.unsa.ba

Food Science in Sarajevo. Bosnian common bean germplasm, which is currently maintained at the above-mentioned collection, has previously been phenotyped (Grahić *et al.*, 2013). Howerever, an integral approach to germplasm evaluation nowadays requires the use of molecular markers, therefore, this study aimed to analyze the genetic diversity of the common bean accessions held at the Gene-bank of the Faculty of Agriculture and Food Science in Sarajevo, using SSR markers. The obtained data will be used to develop a strategy for the conservation and utilization of this genetic resource.

Food Science in Sarajevo. Bosnian common bean germplasm, which is currently maintained at the above-mentioned collection, has previously been phenotyped (Grahić *et al.*, 2013). Howerever, an integral approach to germplasm evaluation nowadays requires the use of molecular markers, therefore, this study aimed to analyze the genetic diversity of the common bean accessions held at the Gene-bank of the Faculty of Agriculture and Food Science in Sarajevo, using SSR markers. The obtained data will be used to develop a strategy for the conservation and utilization of this genetic resource.

## MATERIALS AND METHODS

To obtain plant material for analysis, seeds (3-5) of 21 bean accession from the collection stored in the Gene-bank of the Faculty of Agriculture and Food in Sarajevo were sown in pots filled with substrate. A single pot was used for individual accession. Genomic DNA was extracted from green leaves of the collected common bean seedlings using a peqGOLD plant DNA kit (Peqlab) according to the manufacturer's instructions. Ten primer pairs were used for SSR amplification (Table 1). PCR analysis was performed in a Veriti TM Thermal Cycler (Applied Biosystems®, Foster City, California, USA), according to the protocol published by Carović-Stanko *et al.* (2017).

Table 1. Microsatellite markers used in this study

| Markers | Forward and reverse PCR primer | Repetitive motifs |
|---|---|---|
| BM143 | F: GGGAAATGAACAGAGGAAA<br>R: ATGTTGGGAACTTTTAGTGTG | $(GA)_{35}$ |
| BM151 | F: CACAACAAGAAAGACCTCCT<br>R: TTATGTATTAGACCACATTACTTCC | $(CT)_{14}$ |
| BM157 | F: ACTTAACAAGGAATAGCCACACA<br>R: GTTAATTGTTTCCAATATCAACCTG | $(GA)_{16}$ |
| BM172 | F: CTGTAGCTCAAACAGGGCACT<br>R: GCAATACCGCCATGAGAGAT | $(GA)_{23}$ |
| BM210 | F: ACCACTGCAATCCTCATCTTTG | $(CT)_{15}$ |

| | | |
|---|---|---|
| | R: CCCTCATCCTCCATTCTTATCG | |
| BMd12 | F: CATCAACAAGGACAGCCTCA<br>R: GCAGCTGGCGGGTAAAACAG | $(AGC)_7$ |
| BMd20 | F: GTTGCCACCGGTGATAATCT<br>R: GTGAGGCAAGAAGCCTTCAA | $(TA)_5$ |
| GATS91 | F: GAGTGCGGAAGCGAGTAGAG<br>R: TCCGTGTTCCTCTGTCTGTG | $(GA)_{17}$ |
| PVag001 | F: CAATCCTCTCTCTCTCATTTCCAATC<br>R: GACCTTGAAGTCGGTGTCGTTT | $(GA)_{11}$ |
| Pvctt001 | F: GAGGGTGTTTCACTATTGTCACTGC<br>R: TTCATGGATGGTGGAGGAACAG | $(CTT)_3(T)_3(CT)_6$ |

PCR amplification was carried out in the total volume of 20 µL, containing 2 pmol of the tailed forward primer, 8 pmol of reverse, 8 pmol of FAM-labeled M13 primer, 1 × PCR buffer, 4 pmol of each dNTP, 0.5 U TaqTM HS DNA Polymerase (Takara Bio Inc.) and 5 ng of genomic DNA.
The average number of alleles per locus (Na), observed heterozygosity ($H_O$), and expected heterozygosity ($H_E$) for each of the 10 microsatellite loci was calculated in SpaGedi v.1.2 (Hardy and Vekemans, 2002). Bayesian model-based cluster procedure within Structure ver. 2.2.3 (Pritchard *et al.*, 2000) was used in order to examine the population structure. K (unknown) RPPs (reconstructed panmictic populations) were computed on individuals testing K (log-likelihood) = 1 - 10 for all accessions. It was assumed that sampled cultivars were from an unknown origin. For each K, ten independent runs were conducted. Tests were based on admixture model where allelic frequencies were correlated and which assumes different Fst values for specific subpopulations in a burn-in period of 200,000 and 500,000 iterations. Structure harvester ver. 0.6.1 application (Earl and von Holdt, 2011), which implements the Evanno method (Evanno *et al.*, 2005), was used to estimate the most probable K value for the analyzed data. After determining K value, the individuals were assigned to specific clusters (Vigouroux *et al.*, 2008) via the run with the maximum likelihood. Additional cluster asignment that was based on the average of membership probability values over ten runs, confirmed the results of the aforementioned approach. The assignment of one cultivar in an RPP was provided by the probability of membership qI chosen at 95 %..
An FCA approach (factorial correspondence analysis) that is based on allele frequencies was implemented using Genetix (Belkhir *et al.*, 2001). All input data for statistical software was prepared with the MADC v. 2.0 computer program (Grahić and Grahić, 2017).

## RESULTS AND DISCUSSION

All ten microsatellite loci used for the genetic characterization of 21 examined accessions of common bean were polymorphic. The analyzed loci were also polymorphic in other similar studies (Carović- Stanko *et al.*, 2017; Burle *et al.*, 2010), which were conducted on a larger number of samples. The average Gene diversity value for all analyzed loci was 0.631 and ranged between 0.315 (BMd20) and 0.831 (BM143), while the average number of alleles (Na) was 4.7. In studies conducted by Gioia *et al.* (2019) and Carović-Stanko *et al.* (2017) the values for Gene diversity were lower (0.453 and 0.572), while Carvalho *et al.* (2020) reported slightly higher values (0.680). However, the average number of alleles found by Carvalho *et al.* (2020), Gioia *et al.* (2019), and Carović-Stanko *et al.* (2017) was higher (10.7, 7.3, and 5.6), most likely due to a larger number of analyzed accession (265, 192 and 183) (Table 2).

Table 2. Number of alleles and genetic diversity for 10 SSR loci

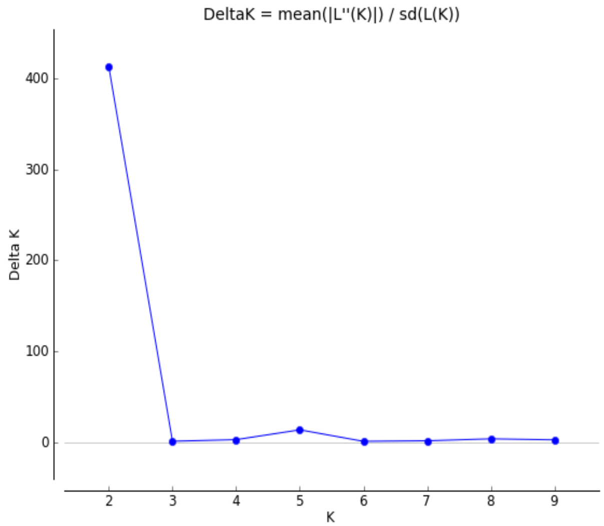
| Analyzed loci | Nuber of alleles (Na) | Gene diversity (Nei, 1987) |
|---|---|---|
| BM157 | 4.0 | 0.6028 |
| BM172 | 4.0 | 0.5621 |
| BM14 | 7.0 | 0.8316 |
| GATS91 | 9.0 | 0.8165 |
| Pvctt001 | 4.0 | 0.5482 |
| BM210 | 5.0 | 0.7096 |
| BMd20 | 2.0 | 0.3159 |
| BMd12 | 3.0 | 0.5540 |
| BM151 | 4.0 | 0.6597 |
| PVag001 | 5.0 | 0.7166 |
| Average | 4.7 | 0.6317 |

The BM172 locus was used to investigate the segregation of the analyzed accessions of B&H common bean according to the gene center of origin. The mentioned locus, according to González Torres *et al.* (2004), is an appropriate tool for preliminary assessment of examined accessions gene center of origin. Namely, based on the alleles detected at this locus it is possible to determinate the phaseolin type of a genotype (80 bp indicates a phaseolin type ‘S’ or Mesoamerican origin, while all other alleles indicate p phaseolin types ‘A’ ‘C’ ‘H’ and ‘T’, with an Andean gene center of origin). In this study, while scoring SSR alleles, an offset of +1 was detected, therefore accessions with allele '81' were classified in the Mesoamerican gene pool. This approach for identifying the gene center of origin has already been used on common bean accessions from B&H (Grahić *et al.*, 2018). Among the 21 analyzed local common bean accessions, 16 belong to the Andean gene pool (76.2%) and only five belong to Mesoamerican (23.8%) (Table 3).

Table 3. Allele variants at the BM172 locus for the analyzed common bean accessions

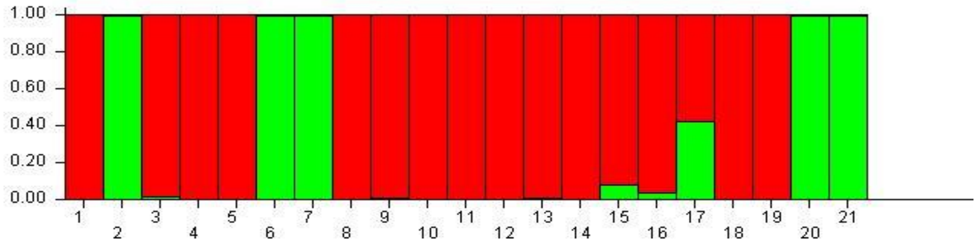
| | | | | | | | |
|---|---|---|---|---|---|---|---|
| BM172 | BH01 | BH02 | BH03 | BH04 | BH05 | BH06 | BH07 |
| | 099:099 | 081:081 | 081:081 | 099:099 | 109:109 | 081:081 | 081:081 |
| BM172 | BH08 | BH09 | BH10 | BH11 | BH12 | BH13 | BH14 |
| | 099:099 | 099:099 | 099:099 | 099:099 | 099:099 | 099:099 | 099:099 |
| BM172 | BH15 | BH16 | BH17 | BH18 | BH19 | BH20 | BH21 |
| | 099:099 | 099:099 | 081:081 | 099:099 | 099:099 | 087:087 | 087:087 |

By analyzing ΔK (Evanno *et al.*, 2005), for 21 Bosnian common bean accessions, the highest value was obtained at K=2 (Chart 1). In the study published by Maras *et al.* (2015), for five common beans germplasms, ΔK analyses revealed maximum values on both K=2 and K=3. Similar results were reported by Carvalho *et al.* (2020) and Carović-Stanko *et al.* (2017).



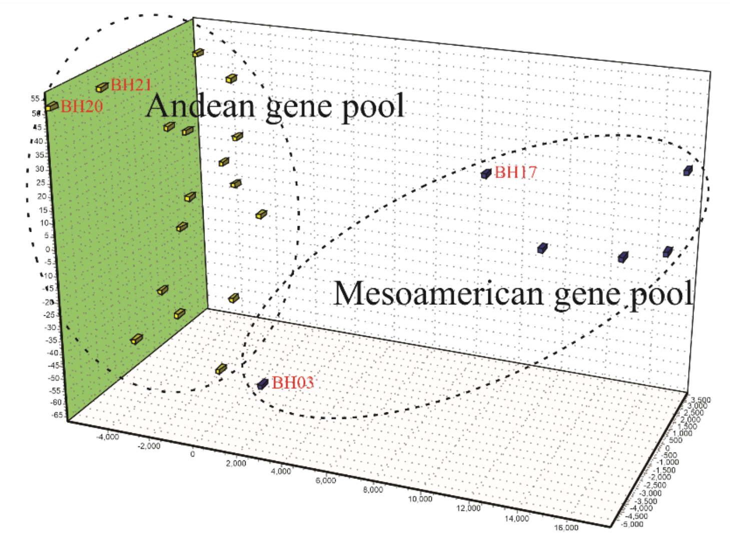

Graph 1. Plot of ΔK values (Evanno *et al.*, 2005) from the Structure analyses of 21 common bean accessions obtained through Structure harvester ver. 0.6.1 (Earl and von Holdt, 2011).

All of the analyzed genotypes were classified into two groups (RPP) with qI > 95%. As opposed to the qI values used by Carović-Stanko *et al.* (2017) and Maras *et al.* (2015) (75% and 80%), we used a higher qI because of the low number of admixed accession. Most of the 21 analyzed accessions were classified into RPP1 (15). The second RPP included 5 accessions, while only one accession had the qI lower than 95% and was not classified in any of the panmictic populations (BH17).



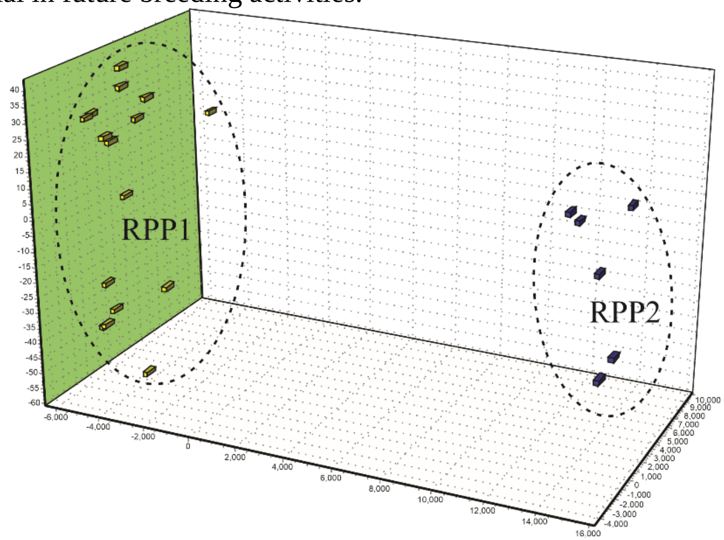
Graph 2. Bar plot of the results from three Bayesian genetic structure analyses of 21 common bean accessions with K = 2.

To get a clearer insight into the genetic relationships between the analyzed groups of genotypes (accessions of Andean and accessions of Mesoamerican origin), an FCA was performed on the molecular data (Graph 3). A certain level of overlap between accession from different gene pools was detected, e.g. accession BH03, which belong to the Mesoamerican gene pool, is very similar to accessions from the Andes. It is also noticeable that larger intragroup divergence manifests between the accessions that are classified into the Mesoamerican gene pool.
Subsequent FCA performed on the RPPs, for K=2, aimed to test the correlation of results obtained with two approaches – grouping based on the gene center of origin, using BM172, and grouping based on the STRUCTURE analysis. Three accessions were classified differently than expected. Accession BH20 and BH21 originating from the Andean gene pool proved to be genetically more similar to Mesoamerican accessions (RPP2). On the other hand, BH03 which possesses the 'S' phaseolin type, was included in the RPP1.
Therefore, Bosnian common bean accessions did not structure completely according to gene center of origin, which was somewhat unexpected. The results obtained are most likely consequences of spontaneous crossing between Mesoamerican and Andean beans that have been cultivated for centuries in B&H. Similar results were reported on the Italian *ex-situ* common bean accessions. Namely, all analyzed accessions were assigned to three different RPPs and the new classification did not match the grouping based on the gene center of origin (Raggi *et al.*, 2012).




Graph 3. Multivariate analysis (FCA) of SSR data for 21 common bean accessions (16 accessions belonging to the Andean gene pool and 5 accession from the Mesoamerican gene pool)

Graph 4 clearly show the differentiation between the reconstructed populations (Graph 4). Looking at the groups individually, it can be noticed that the analyzed common bean accession presents a very heterogeneous genetic material that would, most likely, show great potential in future breeding activities.




Graph 4. Multivariate analysis (FCA) of SSR data for two defined reconstructed populations (RPP) calculated using STRUCTURE (Pritchard *et al.*, 2000) (only genotypes with likelihood of membership to individual RPP above 95% are included in the analyses)

## CONCLUSION

Analyzed accessions of common bean did not completely structure according to the gene center of origin, which points to the fact that some Mesoamerican and Andean common beans present in B&H have been outcrossed during the long period of cultivation. In order to utilize the genetic potential of common bean landraces from B&H, further evaluation studies are needed to asses traits that could be interesting for future breeding programs.

## REFERENCES

Belkhir, K., Borsa, P., Chikli, L., Raufaste, N., Bonhomme, F. (2004): Genetix 4.05, Logiciel Sous Windows TM Pour la Génétique des Populations. Montpellier: Laboratoire Génome, Université Montpellier.

Bitocchi, E., Bellucci, E., Giardini, A., Rau, D., Rodriguez, M., Biagetti, E., Santilocchi, R., Zeuli, P.S., Gioia, T., Logozzo, G., Attene, G., Nanni, L., Papa, R. (2013): Molecular analysis of the parallel domestication of the common bean (*Phaseolus vulgaris*) in Mesoamerica and the Andes. New Phytologist, 197: 300-313.

Burle, M.L., Fonseca, J.R., Kami, J.A., Gepts, P. (2010): Microsatellite diversity and genetic structure among common bean (*Phaseolus vulgaris* L.) landraces in Brazil, a secondary center of diversity. Theoretical and Applied Genetics, 121: 801-813.

Carović-Stanko, K., Liber, Z., Vidak, M., Barešić, A., Grdiša, M., Lazarević, B., Šatović, Z. (2017): Genetic diversity of Croatian common bean landraces. Frontiers in Plant Science. 8: 604.

Carvalho, M.S., de Oliveira Moulin Carias, C.M., Silva, M.A., da Silva Ferreira, M.F., de Souza, T.L.P.O., Posse, S.C.P., Ferreira, A. (2020): Genetic diversity and structure of landrace accessions, elite lineages and cultivars of common bean estimated with SSR and SNP markers. Molecular Biology Reports, 47(9): 6705-6715, doi: 10.1007/s11033-020-05726-7.

Cortés, A.J., Chavarro, M.C., Blair, M.W. (2011): SNP marker diversity in common bean (*Phaseolus vulgaris* L.). Theoretical and Applied Genetics, 123: 827-845.

Earl, D.A., VonHoldt, B.M. (2011): STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources, 4: 359-361.

Evanno, G., Regnaut, S., Goudet, J. (2005): Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology, 14: 2611-2620.

Gestrin, F. (1973): Trgovina i seljačke bune u Sloveniji i Hrvatskoj u 16. stoljeću. Sveučilište u Zagrebu, Institut za hrvatsku povijest, 193-204.

Gioia, T., Logozzo, G., Marzario, S., Spagnoletti Zeuli, P., Gepts, P. (2019): Evolution of SSR diversity from wild types to U.S. advanced cultivars in the Andean and

Mesoamerican domestications of common bean (*Phaseolus vulgaris*). PLoS ONE, 14(1), https://doi.org/10.1371/journal.pone.0211342.g005.

González Torres, R.I., Gaitán, E., Deque, M.C., Toro, O., Debouck, D.G., Thome, J. (2004): Estimation of gene flow on *Phaseolus vulgaris* L. using molecular markers: microsatellites and polymorphisms of chloroplast DNA. 21-25. V Encuentro Latinoamericano y del Caribe de Biotecnología Agrícola, REDBIO, Dominican Republic.

Grahić, A., Grahić, J. (2017): MADC—Marker Analysis Data Compiler user's Manual. (www.divisionagro.ba/apps/docs/madc-marker-analysis-data-compiler/user-manual)

Grahić, J., Ahmetović, A., Gaši, F, Karić, L., Šimon, S., Lazarević, B., Okić, A., Kurtović, M. (2018): Ispitivanje pouzdanosti morfoloških markera pri određivanju gen-centra porijekla običnog graha kolekcionisanog na području Bosne i Hercegovine. Radovi Poljoprivredno-prehrambenog fakulteta Univerziteta u Sarajevu, 68(2): 18-24.

Grahić, J., Gaši, F., Kurtović, M., Karić, L., Đikić, M., Gadžo, D. (2013): Morphological evaluation of common bean diversity in Bosnia and Herzegovina using the discriminant analysis of principal components (DAPC) multivariate method. Genetika, 45(3): 963-977.

Hardy, O.J., Vekemans, X. (2002): SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. Molecular Ecology, 2: 618-620.

Maras, M., Pipan, B., Šuštar-Vozlič, J. (2015): Examination of genetic diversity of common bean from the Western Balkan. Journal of the American Society for Horticultural Science, 140(4): 308-316.

McClean, P., Kamir, J., Gepts, P. (2004): Genomics and genetic diversity in common bean. Legume Crop Genomics, AOCS Press, USA, 60-82.

Papa, R., Nanni, L., Sicard, D., Rau, D., Attene, G. (2006): The evolution of genetic divesity in *Phaseolus vulgaris* L. Chapter 6 in: New approaches to the origins, evolution and conservation of crops, Darwin's harvest, Columbia University Press, USA, 121-142.

Pritchard, J.K., Stephens, M., Donnelly, P. (2000): Inference of population structure using multilocus genotype data. Genetics, University of Oxford, Oxford, United Kingdom, 155(2): 945-959.

Raggi, L., Tiranti, B., Negri, V. (2012): Italian common bean landraces: diversity and population structure. Genetic Resources and Crop Evolution, 60: 1515-1530.

# ISPITIVANJE GENETIČKOG DIVERZITETA LOKALNIH KULTIVARA OBIČNOG GRAHA (*PHASEOLUS VULGARIS* L.) SA PODRUČJA BOSNE I HERCEGOVINE UPOTREBOM MIKROSATELITNIH MARKERA

## Rezime

Istraživanje je imalo za cilj analizirati genetsku raznolikost sorti graha (*Phaseolus vulagris* L.) iz postojeće kolekcije koja se čuva u Gen-banci Poljoprivredno-prehrambenog fakulteta Univerziteta u Sarajevu, uz korištenje mikrosatelitnih markera (SSR). Genotipizirana je ukupno 21 primka graha primjenom deset SSR markera. Dobiveni molekularni podaci su analizirani pomoću analize genetske strukture i faktorske korespondentne analize (FCA). Ispitivane primke graha se nisu u potpunosti grupisale u skladu sa genskim centrima porijekla, što ukazuje na to da je tokom dugog perioda uzgoja graha u Bosni i Hercegovini (BiH) došlo do spontanog ukrštanja mezoameričkih i andskih genotipova. Rezultati analiza potvrđuju da se radi o heterogenom i vrlo zanimljivom genetskom materijalu.

Ključne riječi: *obični grah, mikrosatelitni markeri, gen-centri porijekla, analiza genetičke strukture, faktorijska korespondentna analiza*