DIFFERENT MULTIVARIATE ANALYSIS FOR FRUIT TRAITS IN SWEET PEPPER BREEDING

Dario DANOJEVIĆ* and Slađana MEDIĆ-PAP

Institute of Field and Vegetable Crops, Novi Sad, Serbia

Danojević D. and Medić-Pap (2018): *Different multivariate analysis for fruit traits in sweet pepper breeding.*- Genetika, Vol 50, No.1, 121-129.

Pepper is very popular vegetable crop in Southeast Europe and in Serbia as well. Wide genetic variability is essential in pepper breeding programs. The aim of this study was to evaluate variability for the most important fruit traits and differences between 28 sweet pepper genotypes from the working collection from the Institute of Field and Vegetable Crops (Novi Sad, Serbia). The following traits were analyzed: fruit weight (g), fruit length (cm), fruit diameter (cm), fruit index, number of locules, number of apexes, pericarp thickness (mm), and total soluble solids (°Brix). Results confirmed great variability in evaluated pepper fruit traits. Genotypes were separated into individual groups based on fruit characteristics. According to our research, hierarchical cluster analysis represented differentiation among groups of genotypes more clearly than PCA, but not comparing to k-means. Hierarchical cluster analysis showed the similarity between genotypes, but k-means clustering did not. Genotypes from group 3 will be used in breeding for higher fruit weight and group 6 for thicker pericarp.

Keywords: Capsicum annuum, genotype, fruit characteristic, cluster analysis, principal component analysis

INTRODUCTION

Pepper (*Capsicum* spp.) is one of the world's major vegetable and spice crops (ZEWDIE *et al*., 2004). The genus *Capsicum* consists of 38 species (SILVAR and GARCÍA-GONZÁLEZ 2016). From the 38 species, five is cultivated, and *C. annuum* is the most widely used. Historians believe that pepper has been used in the human nutrition between 7200 and 5200 B.C. in South America (MACNEISH, 1964). Today in the world there are a lot of pepper types and varieties. Pepper (*C. annuum* L.) is one of the major vegetable crops in Serbia. Pepper comparing to other

_____

Corresponding author: Dario Danojević, Institute of Field and Vegetable Crops, Maksima Gorkog 30, 21000 Novi Sad, Serbia; Phone: +381 21 4898 352, Fax: +381 21 4898 355, E-mail: dario.danojevic@ifvcns.ns.ac.rs

vegetable species (without potatoes) has the first rank in Serbia with 16,977 ha in 2016 (STATISTICAL OFFICE OF THE REPUBLIC OF SERBIA, 2017). In different regions of Serbia, consumers are preferred to use pepper fruits with various shapes, sizes and colors (DANOJEVIĆ *et al*., 2016). In Serbia, beside of fresh use, pepper usually used as roasted, stuffed (with various fillings), pickled, dried, ground, ajvar, pindjur and processed in many different ways (DANOJEVIĆ *et al*., 2017). Plant breeders must have genetically diverse germplasm in their breeding collections, not only for yield but also for other important traits. Fruit characterization is the first step in the description and classification of pepper germplasm for breeding purpose. The application of appropriate statistical methods is a useful tool for the description and genotype classification since it enables plant breeders to identify and select valuable genetic resources in breeding programs (JANKULOVSKA *et al*., 2014). In genetic studies, a frequently used analysis is the hierarchical cluster analysis. However, application of some non-hierarchical cluster analyses, such as the k-means could be of help (BABIC *et al*., 2012). According to OCCHIUTO *et al*. (2014) fruit characteristics are the most influential variables in pepper clustering. Also researched with 67 morphological and physiological traits, BOZOKALFA *et al*. (2009) found that the greatest variation was described with fruit traits. During the selection of peppers very important traits are: fruit weight, fruit length, and fruit diameter because they affect the yield directly (SMITH and BASAVARAJA 2005; BHARADWAJ *et al*., 2007). Among the morphological traits the most important for the farmers, tradespeople and consumers are: fruit length, fruit diameter, fruit weight, edible part of fruit and fruit thickness (TODOROVA and DJINOVIC, 2017). Because the fruit traits are the most important in pepper classification, the aim of this research was a screening of the pepper working collection gene pool for main morphological fruit traits and identifying the best sweet pepper parents for further breeding programs.

MATERIALS AND METHODS

Twenty-eight sweet pepper genotypes (old varieties and breeding lines) were sown in the last week of March in 2014 in a plastic greenhouse. The field trials were conducted at the experimental field of the Institute of Field and Vegetable Crops (Rimski Šančevi), Novi Sad (Serbia). Genotypes were transplanted in two replicates (rows) with 20 plants in each row. The density of plants was 70 x 25 cm. Regular cultural practices were applied throughout the growing season (inter-row cultivation, irrigation, treatment against bacterial leaf spot). Five fruits per replicate were harvested in October at the physiological maturity. The following traits were analyzed: fruit weight (g), fruit length (cm), fruit diameter (cm), fruit index (fruit length/fruit diameter), number of locules, number of apexes, pericarp thickness (mm) and total soluble solids-TSS (°Brix). Data were collected according to Descriptor for *Capsicum* (IPGRI, AVRDC, CATIE, 1995). All traits were assessed on 10 ripe fruits per genotype. TSS was measured by digital refractometer. Software package Statistica for Windows ver. 12, (STATSOFT. INC. 2013) was used for Principal Component Analysis (PCA) and Cluster Analysis (CA). Mean values per genotype were standardized (Mean=0, SD=1) and used for analysis. For the construction of dendrograms were used complete linkage and squared euclidean distances. Principal components have been extracted until the Eigenvalue > 1.

RESULTS AND DISCUSSION

On the basis of the analysis of 280 pepper fruits, fruit weight had the highest range 295.6 g., but fruit index and a number of apexes had the highest Coefficient of Variation (CV)

(Tab.1). Total soluble solids had the lowest CV (15.03%). Although the sample was relatively small (28 genotypes) it was characterized by a relatively high variability. FONSECA *et al.* (2008) noted high variability of fruit traits in pepper landraces also.

Table 1. Descriptive statistic of evaluated fruit traits in 28 sweet pepper genotypes

| Trait	Mean	Minimum	Maximum	Range	CV (%)
Fruit Weight (g)	114.69	8.70	304.30	295.60	41.51
Fruit Length (cm)	10.00	3.20	20.00	16.80	33.80
Fruit Diameter (cm)	6.42	1.50	63.50	62.00	59.93
Fruit Index	1.91	0.24	7.69	7.44	66.61
Number of Locules	2.86	1.00	5.00	4.00	23.36
Number of Apexes	2.41	1.00	7.00	6.00	65.98
Pericarp Thickness (mm)	4.80	1.00	9.00	8.00	29.41
Total Soluble Solids (°Brix)	6.30	3.80	9.80	6.00	15.03

PCA indicated that the first two components explained 73.35% of the total variance (Table 2). Since the first two principal components (PC) were over eigenvalue 1, only those were interpreted. The most important positive traits in the first PC were fruit index and fruit length, while the negative were: fruit weight, fruit diameter, number of locules, number of apexes, pericarp thickness and total soluble solids. Also, ILIĆ *et al.* (2013), found in old Serbian pepper populations that fruit yield per plant, fruit weight, and pericarp thickness were the most important variables in the first PC. In Bulgarian peppers, the most strongly correlating traits with the first axis were: fruit width, fruit weight and pericarp thickness (TSONEV *et al.,* 2017). BOZOKALFA *et al.* (2009) found that in the first two PC, the greatest variation was described with fruit traits (fruit diameter, fruit weight, fruit volume, fruit wall thickness, pedicel length and fruit length).

Table 2. Eigenvalues, the proportion of variability between the original variables and the first two principal
components (PC)

Trait	PC1	PC2
Fruit Weight	**-0.781**	0.393
Fruit Length	0.594	**0.708**
Fruit Diameter	**-0.828**	0.287
Fruit Index	**0.949**	-0.042
Number of Locules	**-0.698**	-0.241
Number of Apexes	**-0.773**	-0.537
Pericarp Thickness	**-0.794**	0.242
Total Soluble Solids	0.473	**-0.478**
Eigenvalue	4.493	1.375
% Total variance	56.168	17.187
Cumulative %	56.168	73.355

The results of PCA biplot showed high diversity in the evaluated collection (Graph 1). If there is a large number of genotypes or low variability among them in PCA, they often cannot

be distinguished well. The most distinct pepper genotype in PCA biplot was Slatka feferona. All genotypes with more apexes were on the lower left side from the biplot center. In this group Z. Rotund and B. Rotund (tomato shaped) and genotype 162 were more distinguished. On the other hand all genotypes with one apex were on the upper right side of biplot. Some genotypes from this group (13, 267, Venčarka, 216, and 73) were more separated. Because of the fact that separation among genotypes evaluated with PCA was not transparent, some other multivariate analysis was done to obtain more clear differences.
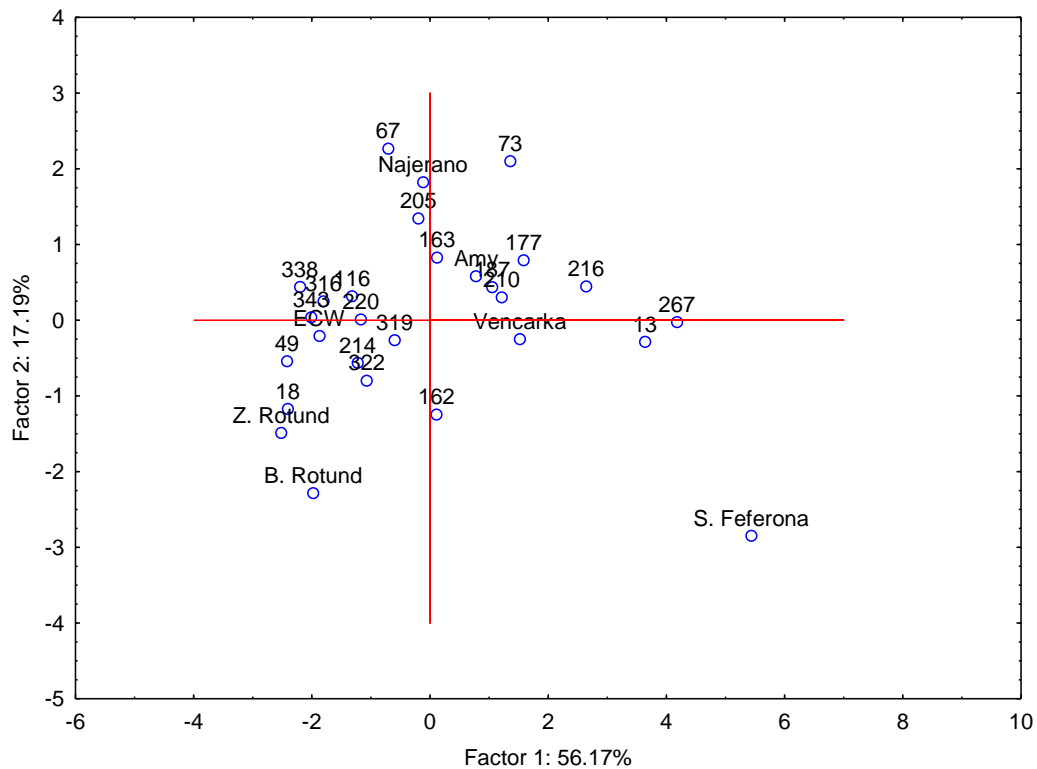
Graph 1. PCA biplot of evaluated sweet pepper genotypes

K-means algorithm has an advantage in comparison to other clustering methods (e.g. hierarchical clustering methods), which have non-linear complexity. Other reasons for the algorithm's popularity are its ease of interpretation, the simplicity of implementation, the speed of convergence and adaptability to sparse data (DHILLON and MODHA, 2001). K-means is not a hierarchical clustering algorithm, but it is a relocation method (THOMAS and HARODE, 2015). Based on the k-means analysis, and eight quantitative measured traits, all 28 genotypes were grouped into 7 clusters. Each cluster had a varied number of genotypes. The cluster 3 consisted of the highest number of genotypes (32.14%). Fruits of cluster 3 had the highest fruit weight and number of locules (Table 3). All genotypes from cluster 3 are bell peppers (Table 4). The second

biggest cluster 4 (25% percent of all genotypes) consisted from mainly kapia type. Tomato shaped peppers (18, Z. Rotund and B. Rotund) were classified in cluster 2. Those genotypes had the lowest fruit length, fruit index, the highest number of apexes and high pericarp thickness. Separation of *C. annuum* genotypes based on fruit shape and fruit characteristics has been previously reported (PARAN *et al*., 1998; GELETA *et al*., 2005; ORTIZ *et al*., 2010; NSABIYERA *et al*., 2013; GONZÁLEZ-PÉREZ *et al.*, 2014, and SALEH *et al.,* 2016). In cluster 7 with the highest fruit length and fruit diameter were grouped two genotypes 37 and 73. The results of ZEČEVIĆ *et al*. (2011) indicate that breeders should cross pepper genotypes with high mean values for breeding traits. Slatka feferona was the only variety in cluster 5 with the lowest fruit weight, fruit diameter, and pericarp thickness, but the highest fruit index and total soluble solids.

Table 3. Mean values of eight pepper fruit traits, number of genotypes and genotype percentage of seven clusters obtained by the K-means method

Cluster	Fruit weight (g)	Fruit Length (cm)	Fruit Diameter (cm)	Fruit Index	Number of Locules	Number of Apexes	Pericarp Thickness (mm)	TSS* (°Brix)	Number of Genotypes	Genotype Percentage (%)
1	53.64	11.97	3.96	3.32	2.30	1.00	3.21	6.32	4	14.29
2	101.94	4.23	7.47	0.57	2.97	5.37	5.77	5.99	3	10.71
3	146.14	9.12	7.22	1.30	3.33	3.29	4.98	6.29	9	32.14
4	115.17	12.06	5.88	2.07	2.74	1.30	4.83	6.05	7	25.00
5	13.97	9.70	1.79	5.42	2.50	1.00	1.62	8.61	1	3.57
6	140.01	6.55	7.56	0.86	2.75	2.85	6.66	6.30	2	7.14
7	137.79	15.13	9.18	1.82	2.35	1.00	5.26	6.41	2	7.14

TSS*-Total Soluble Solids

Table 4. Cluster members of evaluated sweet pepper genotypes obtained by the K-means method
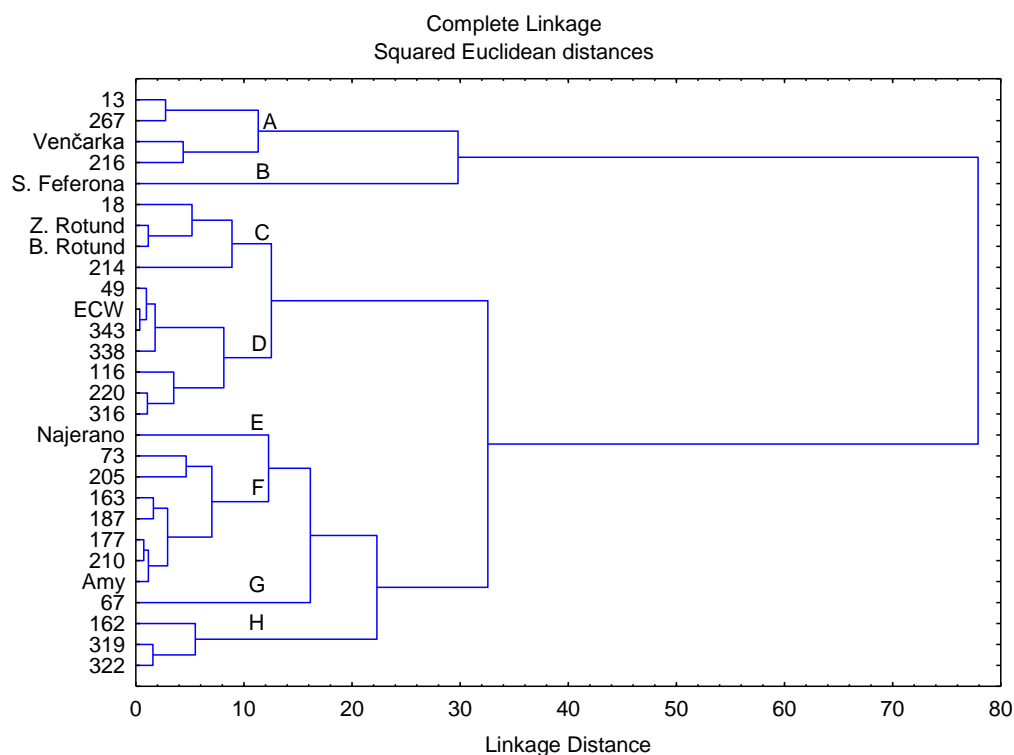
Cluster	Genotype
1	13, Venčarka, 216, 267
2	18, Z. Rotund, B. Rotund
3	49, ECW, 116, 162, 220, 316, 319, 322, 343
4	Najerano, 163, 177, 187, 205, 210, Amy
5	Slatka Feferona
6	214, 338
7	67, 73

By the use of the k-means method, genotypes were clearly divided into different groups, but similarities within the groups were not visible. PCA analysis has a disadvantage, when a large number of genotypes are included. In that case, the genotypes cannot be visually distinguished in a biplot. Hierarchical cluster analysis shows more or less similarity and dissimilarity between genotypes within the same group. KAUR and KAUR, (2013) described that hierarchical clustering shows more quality compared to k-means clustering, but k-means clustering is good for large dataset and hierarchical is good for small datasets.

Complete-link clustering is a method that consider the distance between two clusters to be equal to the longest distance from any member of one cluster to any member of the other cluster (KING, 1967). The complete linkage clustering method usually produces more compact

clusters and more useful hierarchies than the single-link clustering method. But where is the optimal level to cut dendrogram and to determinate optimum number of clusters? One approach is to study the amount of change in the dendrogram, and where changes are the biggest there is the optimum cutting level.

The cutting level was on 11 linkage distance in our research, so there were 8 groups, and pepper genotypes with similar fruit characteristics clustered together (Graph 2). Cluster A in dendrogram had same genotypes as cluster 1 in k-means (13, Venčarka, 216, 267), also S. feferona (cluster B) is alone as in k-means (group 5). Tomato shaped peppers (cluster C) had one more genotype (genotype 214) than group 2. Cluster D consist of bell peppers. Najerano (group E) was distinguished from other genotypes with one apex because of lower total soluble solids, while the genotype 67 (group G) had a maximum fruit weight.

Graph 2. Dendrogram of evaluated sweet pepper genotypes

.

CONCLUSION

Obtained results confirmed great variability in evaluated pepper fruit traits. Genotypes were separated into individual groups based on fruit characteristics. Most groups (5 out of 7) had fruit weight higher than 100 g. Genotypes from group 3 will be used in breeding for higher fruit weight. The genotypes from group 6 will be used in breeding for thicker pericarp. Slatka

feferona was characterized with the highest total soluble solids and it will be used in breeding for this trait. The basis of this research, hierarchical cluster analysis produce visually more clearly groups of genotypes than PCA, but not comparing to k-means. Also, hierarchical cluster analysis shows the similarity between genotypes, but k-means clustering does not. Because of the lower number of genotypes in the trial, hierarchical cluster analysis was more appropriate than PCA and k-means clustering. Evaluating of other fruit traits and comparing genotypes within the same group will be done in future. If similar genotypes from the same group, do not have a difference in the most important fruit traits, the best genotypes will be retained in the working collection only. This research will be continued with other pepper genotypes.

REFERENCES

BABIC V., J. VANCETOVIC, S. PRODANOVIC, V. ANDJELKOVIC, M. BABIC, N. KRAVIC (2012): The Identification of Drought Tolerant Maize Accessions by Two-Step Cluster Analysis. Rom. Agric. Res., *29*: 53- 61.

BHARADWAJ D. N., H. SINGH, R. K. YADAV (2007): Genetic variability and association of component characters for yield in chilli (*Capsicum annuum* L.). Prog. Agric., *7* (1-2): 72-74.

BOZOKALFA M. K., D. ESIYOK, K. TURHAN (2009): Patterns of phenotypic variation in a germplasm collection of pepper (*Capsicum annuum* L.) from Turkey. Span. J. of Agric. Res. *7*: 83-95.

DANOJEVIĆ D., S. MEDIĆ-PAP, A. SAVIĆ, J. ČERVENSKI (2016): Fruit Traits of Pepper Genotypes Originating from Open Pollination. Ratar. Povrt. *53* (2): 69-73.

DANOJEVIĆ D., S. MEDIĆ-PAP, F. FRANETA, M. IGNJATOV, A. TAKAČ, J. ČERVENSKI (2017): Značajni momenti u proizvodnji paprike. Zbornik referata 51. Savetovanje agronoma i poljoprivrednika Srbije (SAPS), Zlatibor, 22-28. Januar 2017. Institut za ratarstvo i povrtarstvo, Novi Sad: 62-70.

DHILLON I. S., D. S. MODHA (2001): Concept Decompositions for Large Sparse Text Data Using Clustering. Mach. Learn. *42*: 143–175.

FONSECA R. M., R. LOPES, W. S. BARROS, M. T. G. LOPES, F. M. FERREIRA (2008): Morphologic Characterization and Genetic Diversity of *Capsicum chinense* Jacq. Accessions along the Upper Rio Negro-Amazonas. Crop Breed. Appl. Biotechnol., *8*: 187-194.

GELETA L. F., M. T. LABUSCHAGNE, C. D. VILJOEN (2005): Genetic Variability in Pepper (*Capsicum annuum* L.) Estimated by Morphological Data and Amplified Fragment Length Polymorphism Markers. Biodivers. Conserv., *14*: 2361-2375. http://dx.doi.org/10.1007/s10531-004-1669-9

GONZÁLEZ-PÉREZ S., A. GARCÉS-CLAVER, C. MALLOR, L. E. S. DE MIERA, O. FAYOS, F. POMAR, C. SILVAR (2014): New Insights into *Capsicum* spp. Relatedness and the Diversification Process of *Capsicum annuum* in Spain. PloS ONE, *9*, e116276. http://dx.doi.org/10.1371/journal.pone.0116276

ILIĆ Z., L. MILENKOVIĆ, M. VASIĆ, Z. GIREK, M. ZDRAVKOVIĆ, J. ZDRAVKOVIĆ (2013): Old Cultivars and Populations From Traditional Pepper-Growing Regions of Serbia as Breeding Potential. J. Agric. Sci., *5* (6): 132-140.

IPGRI, AVRDC, CATIE (1995): Descriptors for Capsicum (*Capsicum* spp.). Rome, Italy: International Plant Genetic Resources Institute; Taipei, Taiwan: Asian Vegetable Research and Development Center; Turrialba, Costa Rica: Centro Agronómico Tropical de Investigación y Enseñanza.

JANKULOVSKA M., S. IVANOVSKA, A. MARJANOVIC-JEROMELA, S. BOLARIC, LJ. JANKULOSKI, Z. DIMOV, D. BOSEV, B. KUZMANOVSKA (2014): Multivariate analysis of quantitative traits can effectively classify rapeseed germplasm. Genetika, *46* (2): 545-559.

KAUR M., U. KAUR (2013): Comparison Between K-Mean and Hierarchical Algorithm Using Query Redirection. Int. J. Adv. Res. Comput. Sci. Softw. Eng., *3* (7): 1454-1459.

KING B. (1967): Step-wise Clustering Procedures, J. Am. Stat. Assoc. *6*9: 86-101.

MACNEISH R. S. (1964): Ancient Mesoamerican civilization. Science. *143*: 531-537.

NSABIYERA V., M. L. M. OCHWO-SSEMAKULA, P. SSERUWAGI, P. GIBSON, C. OJIEWO (2013): Morphological Characterization of Local and Exotic Hot Pepper (*Capsicum annuum* L.) Collections in Uganda. Biorem. Biodiv. Bioavail., *7*: 22-32.

OCCHIUTO P. N., I. E. PERALTA, P. D. ASPRELLI, C. R. GALMARINI (2014): Characterization of Capsicum Germplasm Collected in Northwestern Argentina Based on Morphological and Quality Traits. AgriScientia, *31*: 63-73.

ORTIZ R., F. DELGADO DE LA FLORA, G. ALVARADO, J. CROSSA (2010): Classifying vegetable genetic resources—A case study with domesticated *Capsicum* spp. Sci. Hort. *126*: 186–191.

PARAN I., E. AFTERGOOT, C. SHIFRISS (1998): Variation in *Capsicum annuum* revealed by RAPD and AFLP markers. Euphytica *99*: 167–173.

SALEH B. K., R.W. KASILI, E. G. MAMATI, W. ARAIA, A. B. NYENDE (2016): Classification of Local Pepper Collections (*Capsicum* spp.) from Eritrea Using Morphological Traits. Am. J. Plant. Sci., *7*: 590-600.

SILVAR C., C. A. GARCÍA-GONZÁLEZ (2016): Deciphering Genetic Diversity in the Origins of Pepper (*Capsicum* spp.) and Comparison with Worldwide Variability. Crop Sci., *56*: 3100-3111.

SMITH R. P., N. BASAVARAJA (2005): Variability and correlation studies in chilli (*Capsicum annuum* L.). Karnataka J. Agric. Sci., *19* (4): 888-891.

STATISTICAL OFFICE OF THE REPUBLIC OF SERBIA (2017): Database. RZS, Belgrade (http://webrzs.stat.gov.rs/WebSite/public/ReportView.aspx).

THOMAS S. T., U. HARODE (2015): A comparative study on k-means and hierarchical clustering. Int. J. Elect. Comput. Syst. *4* (2): 5-11.

TODOROVA V., I. DJINOVIC (2017): Assessment of Serbian pepper varieties grown in conditions of south Bulgaria. Genetika, *49* (1): 161-172.

TSONEV S., V. TODOROVA, S. GROSEVA, T. POPOVA, E.G. TODOROVSKA (2017): Evaluation of diversity in Bulgarian pepper cultivars by agronomical traits and ISSR markers. Genetika, *49* (2): 647-662.

ZEČEVIĆ B., R. ĐORĐEVIĆ, A. BALKAYA, J. DAMNJANOVIĆ, M. ĐORĐEVIĆ, A. VUJOŠEVIĆ (2011): Influence of parental germplasm for fruit characters in F1, F2 and F3 generations of pepper (*Capsicum annuum* L.). Genetika, *43* (2): 209-216.

ZEWDIE T., N. K. TONG, P. W. BOSLAND (2004): Establishing core collection of *Capsicum* using a cluster analysis with enlightened selection of accessions. Genet. Resour. Crop Evol., *51*: 147-151. doi:10.1023/B:GRES.0000020858.96226.38.

RAZLIČITI METODI MULTIVARIJACIONE ANALIZE ZA SVOJSTVA PLODA SLATKE PAPRIKE

Dario DANOJEVIĆ*, Slađana MEDIĆ-PAP
Institut za ratarstvo i povrtarstvo, Novi Sad, Srbija

Izvod

Paprika je veoma popularna povrtarska vrsta u Srbiji. Široka genetička varijabilnost je od suštinskog značaja u oplemenjivanju paprike. Cilj ovog istraživanja je bio da se ustanovi varijabilnost najvažnijih svojstava plodova između 28 genotipova slatke paprike iz radne kolekcije Instituta za ratarstvo i povrtarstvo (Novi Sad, Srbija). Analizirana su sledeća svojstva: masa ploda (g), dužina ploda (cm), prečnik ploda (cm), indeks ploda, broj komora, broj vrhova, debljina perikarpa (mm) i sadržaj suve materije (°Brix). Rezultati su potvrdili veliku varijabilnost u ispitivanim osobinama ploda. Genotipovi su izdvojeni u pojedinačne grupe. Prema našem istraživanju, hijerarhijska klaster analiza je ustanovila razlike između grupa genotipova jasnije nego metoda glavnih komponenti, ali ne u odnosu na metod k-sredina. Hijerarhijska klaster analiza je pokazala i sličnost između genotipova, ali grupisanje na osnovu k-sredina nije. Genotipovi iz grupe 3 će se koristiti u oplemenjivanju za veću masu ploda, a iz grupe 6 za deblji perikarp.