

# Classification of rapeseed genotypes based on quantitative traits using multivariate techniques

Mirjana Jankulovska<sup>1\*</sup> · Dragana Rajković<sup>2</sup> · Dragana Miladinović<sup>2</sup> · Sreten Terzić<sup>2</sup> ·  
Nada Grahovac<sup>2</sup> · Ankica Kondić-Špika<sup>2</sup> · Ana Marjanović Jeromela<sup>2</sup>

<sup>1</sup>Ss. Cyril and Methodius University in Skopje, Faculty of Agricultural Sciences and Food, Skopje, Republic of North Macedonia

<sup>2</sup>Institute of Field and Vegetable Crops, Novi Sad, Republic of Serbia

**Summary:** The genetic variability of 13 rapeseed genotypes created in Serbia was assessed at the Institute of Field and Vegetable Crops, Novi Sad, in three consecutive years. The genotypes were evaluated and classified based on 10 quantitative traits (plant height, height to the first branch, stem width, number of lateral branches, number of leaves, number of pods, seed yield per plant, 1000 seed weight, oil content and protein content), using principal component analysis (PCA) and two-way cluster analysis. The first four PCs with eigenvalues >1 contributed to 83% of the total explained variance. The PC analysis revealed that plant height, height to the first branch, number of lateral branches and number of leaves contributed maximum to the total divergence of the material. The genotypes were classified in two main clusters: the first cluster comprised three genotypes, while the other genotypes belonged to the second cluster which can further be divided to two subclusters. One genotype was not classified in these two clusters. The selected genotypes that have desirable traits may be used in hybridization programs for improvement of seed yield and traits that contribute to seed yield most.

**Keywords:** *Brassica napus*, PCA, quantitative traits, two-way cluster analysis

## Introduction

Aiming to ensure efficient rapeseed production, rapeseed breeders created varieties that combined high yield, high oil content and specific quality traits (low levels of erucic acid in oil and glucosinolates in seeds). Consequently, it resulted with narrowed genetic basis of elite breeding material and limited variability regarding many characters of value for breeders (Cowling, 2007; Ananga et al., 2008; Kozak et al., 2011). Furthermore,

there are no naturally occurring rapeseed wild forms that can be utilized as variability sources (Lopatinska et al., 2021). The extended knowledge about the diversity that affects specific trait if interest is a prerequisite for their improvement through selection, which has been successful in many crops. By using divergent parental forms, the breeder can achieve improvement of traits of interest (Lopatinska et al., 2021).

Various techniques for evaluation of genetic variation in *Brassica napus*, that could help breeders and geneticists to understand the structure of their germplasm and to predict combinations that would produce the best offspring, are available, such as morphological, biochemical and molecular markers (Marjanovic-Jeromela et al., 2009). The first step in the description and classification of germplasm continues to be its morphological characterization. Seed and oil yield, as ultimate goals in rapeseed breeding programs are complex traits, with highly plastic structure (Diepenbrock, 2000). Yield components have significant effect on the final yield performance. Consequently, the indirect selection for these traits can be effective approach in a rapeseed breeding program (Sadat et al., 2010; Marjanovic-Jeromela et al., 2011). The use of appropriate statistical methods to analyze the data is essential for initial

Corresponding author:  
mirjanajankulovska@yahoo.com

Acknowledgements: This work was done as a part of activities of Centre of Excellence for Innovations in Breeding of Climate Resilient Crops - Climate Crops, Institute of Field and Vegetable Crops, Novi Sad, Serbia. This research was financially supported by the Ministry of Education, Science and Technological Development, Republic of Serbia (Project TR31025 and 451-03-9/2021-14/200032).

Cite this article:

Jankulovska M., Rajković D., Miladinović D., Terzić S., Grahovac N., Kondić-Špika A., Marjanović Jeromela A. (2021). Classification of rapeseed genotypes based on quantitative traits using multivariate techniques. *Ratar. Povrt.*, 58 (3), 88-93.



Copyright © The Author(s) 2021

description and classification of genotypes, as well as for understanding yield constraints.

For increasing the genetic variability and developing rapeseed cultivars with high seed and oil yield potential, exploitation of the existing gene pool is of foremost importance. Different multivariate techniques have been used to characterize the diversity in different rapeseed collections (Asghari et al., 2011, Jankulovska et al., 2014; Sharafi et al., 2015; Kumar & Singh, 2018; Parvin et al., 2019). The objectives of this study were to assess the extend and patterns of variability within the created rapeseed breeding material and to identify quantitative traits which explain most of the variability of analyzed genotypes.

## Material and Methods

The genetic variability of 13 winter type rapeseed genotypes (Sremica, Banaćanka, UM-1, UM-2, UM-5, UM-6, UM-8, UM-9, UM-10, UM-11, UM-12, UM-13 and UM-14) created in Serbia was assessed during three consecutive years at locality Rimski Šančevi (45°19'54.0"N 19°50'12.2"E). The experimental layout was randomized block design with three replications. Each genotype was sown in 4 rows with 25 cm distance between rows and 5 cm within rows. Standard cultivation practices were applied during the vegetative growth.

The genotypes were evaluated for 10 quantitative traits (plant height-PH, height to the first branch-HFB, stem width-SW, number of lateral branches-NB, number of leaves-NL, number of siliques-NP, seed weight per plant-SWP, 1000 seed weight-W1000, oil content-OC and protein content-PC). For each genotype 33 rapeseed plants were measured. Plant height was measured from the root neck to the highest bud on the central branch, and the height of the first side branch from the root neck to the first branch. The thickness of the main stem was measured with a vernier caliper. The number of branches was counted, while the number of leaves was determined by counting the scars

after their fall. Silique number represented both full and empty siliques. To obtain seed yield per plant, the seed weight of each individual plant was measured. Thousand seed weight was calculated from the weight of three samples of 200 seeds for each replication. The oil and protein content were determined by NMR and Kjeldahl method, respectively.

In order to see the differences in the analyzed traits between the genotypes included in this study, as well as their mutual association, descriptive statistics, correlation coefficients and ANOVA with post-hoc Duncan test were applied. Genotype classification was carried out based on standardized values (mean=0, sd=1) for the traits by principal component analysis (PCA) and two-way cluster analysis, using R3.5.1 statistical software.

## Results and discussion

Summarized features of analyzed quantitative traits are presented within Table 1. During the years, plant height varied in range 69.50-163.50 cm. The minimum observed value for number of silique was 62.70, and the maximum value was 261.40. Seed weight per plant had the highest and oil content the lowest coefficient of variation.

Plant height was in middle strong positive correlation with the number of branches and the seed weight per plant. Higher plants imply more branches and consequently more siliques and seeds, thus selection of higher plants may influence higher yield (Bennett et al., 2017; Khan et al., 2017). However, higher plants in general are prone to lodging if lignin and cellulose content, as well as stem width don't increase (Zhang et al., 2017). Middle strong correlations were also detected between the silique number and the protein content (positive correlation) on one hand and the oil and protein content (negatively correlated). Number of leaves was in strong positive correlation with height to the first branch.

Table 1. Descriptive statistics of examined traits

Trait	Unit	Minimum	Maximum	Mean	Std. Error	CV (%)
PH	cm	69.50	163.50	123.06	2.59	22.77
HFB	cm	18.80	88.30	56.41	1.91	36.65
SW	g	0.75	1.46	1.03	0.01	15.36
NB	-	3.60	10.80	6.68	0.12	19.26
NP	-	62.70	261.40	124.92	3.47	30.08
NL	-	6.30	15.70	10.75	0.22	21.67
PC	%	16.87	25.36	20.29	0.19	10.24
OC	%	36.23	49.72	44.24	0.32	7.87
W1000	g	2.97	4.50	3.79	0.03	9.47
SWP	g	1.58	19.81	8.31	0.40	51.58

Std error – standard error, CV – coefficient of variation

Table 2. Phenotypic correlations among 10 quantitative traits in 13 examined rapeseed genotypes over three years

	PH	HFB	SW	NB	NP	NL	PC	OC	W1000
HFB	0.51								
SW	0.11	-0.07							
NB	0.57*	0.38	-0.15						
NP	0.14	0.11	0.42	-0.27					
NL	0.25	0.78**	-0.06	0.48	0.04				
PC	-0.28	0.26	0.14	-0.14	0.53*	0.14			
OC	0.45	0.06	0.18	0.08	-0.08	0.07	-0.69*		
W1000	-0.03	0.03	-0.47	-0.06	-0.25	-0.05	-0.11	-0.05	
SWP	0.57*	-0.09	0.48	-0.07	0.30	-0.51	-0.08	0.22	-0.08

\* Correlation is significant at the 0.05 level

\*\* Correlation is significant at the 0.01 level

Regarding mean genotypic values for plant height, UM-5 had the lowest and UM-11 the tallest main stem (Table 3). Mean number of siliques ranged from 104.90 in UM-1 to 170.78 in UM-12. Cultivar Banačanka had the highest seed weight and oil content. Apart from Banačanka, UM-2 and UM-9 also had high oil content and are considered as potential genotypes for further selection in terms of oil content improvement. Thousand seed weight was the lowest in UM-5 and the highest in UM-2.

Principal component analysis was applied in order to determine contribution of analyzed quantitative traits in the overall variation.

The first four PCs with eigenvalues >1 contributed to 83% of the total explained variance. The PC analysis

revealed that plant height, height to the first branch, number of lateral branches and number of leaves contributed maximum to the total divergence of the studied material, because these traits had the highest correlation with PC1 and explained 27.171% of the total variance (Table 4). The second principal component was highly associated with NP and PC, explaining 21.708% of the total variance. Oil content had the highest negative correlation with PC2. The main contributors to the PC3 were SW and OC, while W1000 was negatively associated with the third principal component. Seed weight per plant had the strongest correlation with PC4. The principal component analysis was effective in identification of traits that contributed most to the variation of the analyzed genotypes.

Table 3. Mean values for analyzed traits

Genotype	PH	HFB	SW	NB	NP	NL	PC	OC	W1000	SWP
UM-1	127.01 cd	62.06 e	0.99 ab	6.76 bcde	104.90 a	10.97 cde	19.47 b	45.45 d	3.94 cd	9.32 ef
UM-10	124.59 bc	54.73 bcd	1.01 abc	6.36 abcd	136.62 cd	10.01 abc	20.17 bc	43.08 bc	4.01 d	9.13 f
Sremica	126.60 cd	52.56 b	1.08 cd	6.78 bcde	147.94 d	9.37 a	21.86 e	44.03 c	3.83 c	11.48 g
UM-11	131.50 cd	68.77 e	1.02 bc	8.09 f	125.56 bc	12.53 f	19.81 bc	46.28 d	3.57 b	7.55 bc
UM-12	118.76 ab	61.60 e	1.07 cd	5.81 a	170.78 e	11.87 ef	22.89 f	43.45 bc	3.68 b	6.87 abc
Banačanka	124.80 bc	52.43 b	1.15 e	6.71 bcde	119.99 ab	11.11 cde	18.47 a	46.42 d	3.88 cd	8.87 def
UM-13	119.56 ab	57.21 bcde	0.95 ab	7.10 de	108.03 a	11.44 de	21.16 de	41.97 a	3.96 cd	6.51 ab
UM-14	119.83 ab	53.19 bc	0.94 a	7.34 e	107.90 a	10.99 cde	20.56 cd	42.41 ab	3.96 cd	6.23 a
UM-2	119.46 ab	57.54 bcde	0.97 ab	6.05 ab	108.47 a	10.40 abcd	19.54 b	45.50 d	4.20 e	7.42 abc
UM-5	116.36 a	52.32 b	1.14 e	6.41 abcd	111.96 ab	10.17 abc	21.30 de	43.06 bc	3.41 a	7.96 cde
UM-6	124.70 bc	59.89 de	1.01 abc	6.29 abcd	117.17 ab	10.61 bcd	20.15 bc	43.56 c	3.60 b	8.88 def
UM-8	126.911 cd	58.09 cde	1.10 cd	6.94 cde	142.60 d	10.80 cde	20.04 bc	43.95 c	3.62 b	10.03 f
UM-9	119.71 ab	42.91 a	0.98 ab	6.24 abc	122.00 abc	9.53 ab	18.29 a	45.97 d	3.55 b	7.83 cd
Average	123.06	56.41	1.03	6.68	124.92	10.75	20.29	44.24	3.79	8.31

Mean values with the same lowercase letters belong to the same group according to the post-hoc Duncan test

Table 4. PC scores for the analyzed quantitative traits in rapeseed genotypes

Analyzed traits	PC 1	PC 2	PC 3	PC 4
PH	<b>0.903</b>	-0.205	-0.062	0.103
HFB	<b>0.881</b>	0.195	-0.186	0.001
SW	0.047	0.078	<b>0.873</b>	0.268
NB	<b>0.788</b>	-0.034	0.260	-0.393
NP	-0.032	<b>0.753</b>	0.387	0.261
NL	<b>0.680</b>	0.068	-0.043	-0.585
PC	0.096	<b>0.965</b>	-0.005	0.003
OC	0.092	<b>-0.757</b>	<b>0.288</b>	0.215
W1000	0.119	0.088	<b>-0.824</b>	<b>0.168</b>
SWP	-0.075	0.004	0.064	<b>0.950</b>
eigenvalue	3.039	2.176	1.818	1.268
% of variance explained	27.171	21.708	17.870	16.249
Cumulative variance	27.171	48.879	66.749	82.998

Two-way clustering of genotypes and traits allows an effective overview and enables an easy determination of the traits based on which the breeding material has been grouped (Jankulovska et al., 2014). As explained by Hageman et al. (2012), this technique is able to extract relevant features from the data set.

Based on the analyzed traits, the genotypes can be classified in two main clusters (Figure 1). The first cluster comprised three genotypes (Banačanka, UM-5 and UM-9), characterized by high values for SWP, SW and OC and low to medium high values for HFB, NB and PH. The other genotypes belonged to the second cluster which can further be divided into two subclusters. The genotypes with the highest seed yield per plant belonged to this cluster (Sremica, UM-10, UM-1, UM-8 and UM-6). One genotype (UM-11) was not classified in these two clusters. This genotype is very valuable, since it had the highest average values for 4 traits (PH, HFB, NB and NL) and the second highest value for OC.

Many researchers applied multivariate statistical methods in order to assess the level of divergence

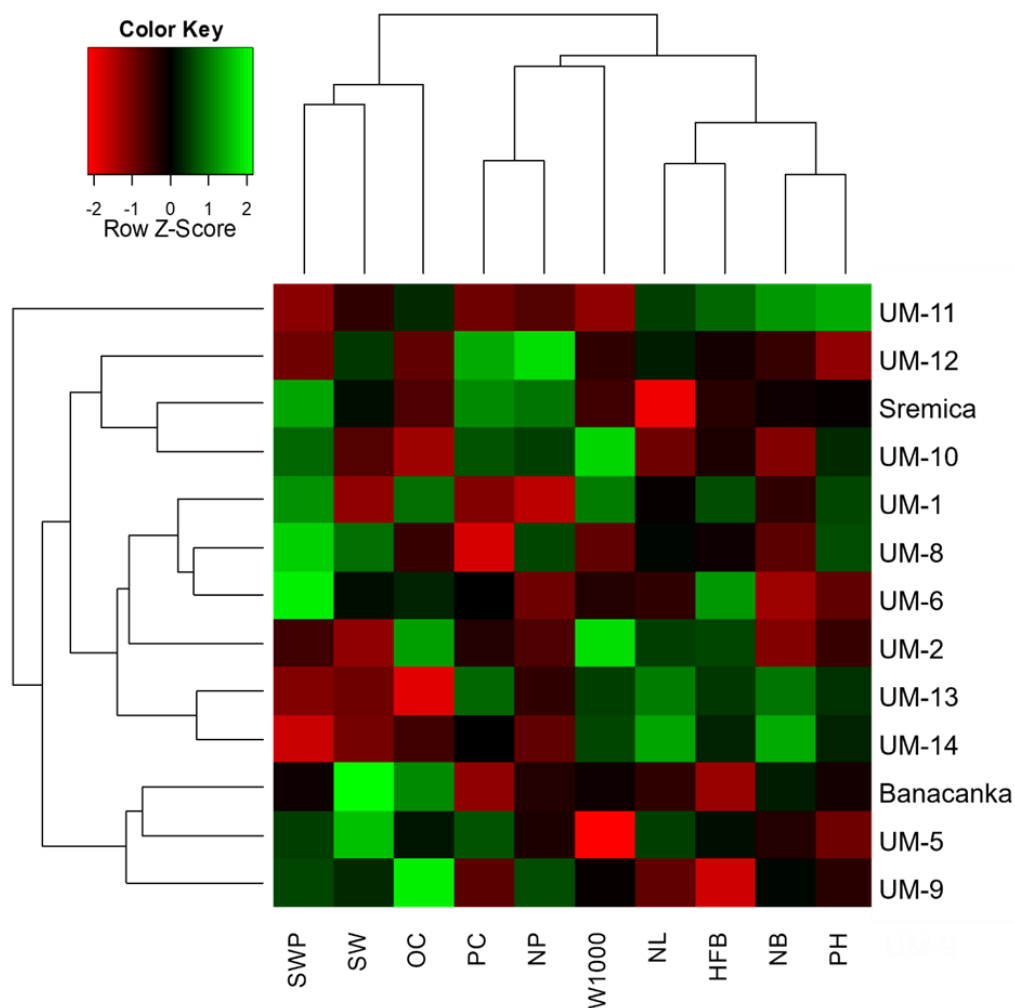


Figure 1. Two-way clustering of genotypes and analyzed traits

within different rapeseed germplasm by analyzing desirable traits for improvement of seed and oil yield, and identified various traits and diverse genetic variability within analyzed populations (Jankulovska et al., 2014; Sharafi et al., 2015; Parvin et al., 2019).

The identified genotypes with a combination of desirable traits will be used in hybridization programs for direct or indirect improvement of seed and oil yield in rapeseed.

## Conclusions

The analyzed rapeseed genotypes significantly differed in most of the analyzed traits, which serves as a good basis for further improvement. Both PC analysis and two-way cluster analysis were successful in identification of traits that contributed the most to the variability of the germplasm, as well as the genotypes with a combination of desirable traits which can be included in future rapeseed breeding programs.

## References

- Ananga A.O., Cebert E., Soliman K., Kantety R., Konan K. & Ochieng J.W. (2008). Phylogenetic relationships within and among Brassica species from RAPD loci associated with blackleg resistance. *Afr. J. Biotechnol.* 7(9): 1287-1293.
- Asghari A., Shokrpour M., Chamanabad H.M. & Sofalian O. (2011). Evaluating genetic diversity of canola cultivars using morphological traits and molecular markers. *Romanian Biotechnological Letters* 16(4):6305-6312.
- Bennett E.J., Brignell C.J., Carion P.W.C., Cook S.M., Eastmond P.J., Teakle G.R., Hammond J.P., Love C., King G.J., Roberts J.A. & Wagstaff C. (2017). Development of a statistical crop model to explain the relationship between seed yield and phenotypic diversity within the *Brassica napus* gene pool. *Agronomy* 7(2):31. <https://doi.org/10.3390/agronomy7020031>
- Cowling A.W. (2007). Genetic diversity in Australian canola and implications for crop breeding for changing future environments. *Field Crop Res.* 104: 103-111.
- Diepenbrock W. (2000). Yield analysis of winter oilseed rape (*Brassica napus* L.): a review. *Field Crop Res.* 67: 35-49.
- Hageman J.A., Malosetti M. & van Eeuwijk F.A. (2012). Two-mode clustering of genotype by trait and genotype by environment data. *Euphytica* 183, 349–359 <https://doi.org/10.1007/s10681-010-0236-6>
- Jankulovska M., Ivanovska S., Marjanovic-Jeromela A., Bolaric S., Jankuloski L.J., Dimov Z., Bosev D. & Kuzmanovska B. (2014). Multivariate analysis of quantitative traits can effectively classify rapeseed germplasm. *Genetika*, 46(2), 545-559.
- Lopatynska A., Bocianowski J., Cyplik A. & Wolko J. (2021). Multidimensional Analysis of Diversity in DH Lines and Hybrids of Winter Oilseed Rape (*Brassica napus* L.). *Agronomy* 11, 645. <https://doi.org/10.3390/agronomy11040645>
- Khan, S., Anwar, S., Kuai, J., Ullah, S., Fahad, S., & Zhou, G. (2017). Optimization of Nitrogen Rate and Planting Density for Improving Yield, Nitrogen Use Efficiency, and Lodging Resistance in Oilseed Rape. *Frontiers In Plant Science*, 8, 532. <https://doi.org/10.3389/fpls.2017.00532>
- Kozak M., Bocianowski J., Liersch A., Tartanus M., Bartkowiak-Broda I., Piotto F.A. & Azevedo, R.A. (2011). Genetic divergence is not the same as phenotypic divergence. *Mol. Breed.* 28, 277–280
- Kumar M. & Singh R.K. (2018). Diversity analysis in parental lines of winter oilseed rape (*Brassica napus* L.). *J. Oilseeds Res.* 35, 94–97.
- Marjanovic-Jeromela A., Kondic-Spika A., Saftic-Pankovic D., Marinkovic R. & Hristov N. (2009). Phenotypic and molecular evaluation of genetic diversity of rapeseed (*Brassica napus* L.) genotypes. *African Journal of Biotechnology* 8 (19): 4835-4844.
- Marjanovic-Jeromela A., Marinković R., Ivanovska S., Jankulovska M., Mijić A. & Hristov N. (2011). Variability of yield determining components in winter rapeseed (*Brassica napus* L.) and their correlation with seed yield. *Genetika*, 43 (1): 51 -66.
- Parvin E., Mahmud F., Bhuiyan S.R. & Haque M. (2019). Multivariate Analysis of Genetic Variation in Rapeseed (*Brassica napus* L.). *Agric. Food Sci. Res.* 6, 1–8.
- Sadat H.A., Nematzadeh G., Jelodar N.B. & Chapi O.G. (2010). Genetic evaluation of yield and yield components at advanced generations in rapeseed (*Brassica napus* L.). *African Journal of Agricultural Research*, 5, 1958-1964.
- Sharafi Y., Majidi M.M., Jafarzadeh M. & Mirlohi A. (2015). Multivariate Analysis of Genetic Variation in Winter Rapeseed (*Brassica napus* L.) Cultivars. *J. Agric. Sci. Technol.* 17, 1319–1331
- Zhang M., Wang H., Yi Y., Ding J., Zhu M., Li C., ... Zhu X. (2017). Effect of nitrogen levels and nitrogen ratios on lodging resistance and yield potential of winter wheat (*Triticum aestivum* L.). *PLOS ONE*, 12(11), e0187543. <https://doi.org/10.1371/journal.pone.0187543>

## Klasifikacija genotipova uljane repice zasnovana na kvantitativnim osobinama primenom multivarijacione analize

Mirjana Jankulovska · Dragana Rajković · Dragana Miladinović · Sreten Terzić ·  
Nada Grahovac · Ankica Kondić-Špika · Ana Marjanović Jeromela

**Sažetak:** Genetička varijabilnost 13 genotipova uljane repice stvorenih u Srbiji, analizirana je u Institutu za ratarstvo i povrtarstvo u Novom Sadu, tokom tri uzastopne godine. Genotipovi su procenjeni i klasifikovani na osnovu 10 kvantitativnih osobina (visina biljke, visina do prve grane, širina stabljike, broj bočnih grana, broj listova, broj ljuski, prinos semena po biljci, masa 1000 semena, sadržaj ulja i sadržaj proteina), koristeći analizu glavnih komponenti i dvosmernu klaster analizu. Prve četiri PC ose doprinele su 83% od ukupne objašnjene varijanse. PC analiza je pokazala da su visina biljke, visina do prve grane, broj bočnih grana i broj listova maksimalno doprineli ukupnoj varijabilnosti materijala. Genotipovi su klasifikovani u dva glavna klastera: prvi klaster se sastojao od tri genotipa, dok su ostali genotipovi pripadali drugom klasteru koji se dalje može podeliti na dva podklastera. Jedan genotip klasifikovan je izvan oba klastera. Odabrani genotipovi koji imaju poželjne osobine mogu se koristiti u programima hibridizacije za poboljšanje prinosa semena i osobina koje najviše doprinose prinosu semena.

**Ključne reči:** *Brassica napus*, dvosmerna klaster analiza, kvantitativne osobine, PCA

This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.