

# Combinative breeding for large seeds in soybean

Galina NAYDENOVA<sup>1,2</sup>, Mariana RADKOVA<sup>3</sup>, Anelia IANTCHEVA<sup>3</sup>

Received January 8, 2020; accepted March 21, 2020.  
Delo je prispelo 8. januarja 2020, sprejeto 21. marca 2020

## Combinative breeding for large seeds in soybean

**Abstract:** Technological qualities of the seeds, including their mass, play an important role in the purposeful use of soybean for food production. The purpose of this study is to determine the potential of specific crosses and recombinant lines in the combinative breeding of high yielding large-seeded soybean varieties. During the period of 2018-2019 the F<sub>3</sub> and F<sub>4</sub> hybride generations of crosses with participation of the ultra-early mature and large-seeded cultivar *Romantica* were studied. Data were used to evaluate: presence and extent of positive transgressive forms by absolute seed mass in F<sub>3</sub> family crosses; genotypic diversity and additive variance at specific crosses; the effectiveness of selection of the trait large seed in F<sub>3</sub>. According to the results, transgressive selection can be successfully used to reach the goal large seeds in soybean. The efficiency of selection of transgressive forms in F<sub>3</sub> generation is high. The genetic potential to combine a high specific mass of seeds with a high yield of seeds per plant has been established for the *Romantica* cross with the Bulgarian standard variety *Srebrina*. Recombinant lines suitable for intensive selection for the trait large seed were obtained from the 'Romantica' x 'Oria' combination. The 'Saikai 20' x 'Romantica' cross possess a very high degree of transgressive segregations.

**Key words:** soybean; combinative breeding; absolute seed mass

## Kombinacijsko žlahtnjenje soje za večja semena

**Izvleček:** Tehnološke lastnosti semen, vključno z njihovo maso imajo pomembno vlogo pri namenski rabi soje v pridelavi hrane. Namen raziskave je bil določiti potencial specifičnih križanj in rekombinantnih linij v kombinirani vzgoji visokodonosnih sort soje z velikimi semeni. V obdobju 2018-2019 so bile preučevane F<sub>3</sub> in F<sub>4</sub> hibridne generacije iz križanj zelo zgodaj dozorevajočih sort soje in sorte *Romantica* z velikimi semeni. Pridobljeni podatki so bili ovrednoteni glede na naslednje parametre: prisotnost in obseg pozitivnih transgresivnih oblik z veliko absolutno maso semen v družini F<sub>3</sub> križancev; genotipsko raznolikost in aditivno spremenljivost pri specifičnih križanjih; učinkovitost izbora lastnosti velikih semen v F<sub>3</sub> generaciji. Glede na rezultate je transgresivna selekcija lahko uspešno uporabljena za vzgojo sort soje v velikimi semeni. Učinkovitost izbora transgresivnih oblik v F<sub>3</sub> generaciji je bila velika. Genetski potencial za kombinacijo lastnosti velika specifična masa semen z velikim pridelkom semen na rastlino je bil dosežen pri križanju sorte *Romantica* s standardno bolgarsko sorto *Srebrina*. Rekombinantne linije, primerne za intenzivno selekcijo za lastnost velika semena so bile dobljene s kombinacijo 'Romantica' x 'Oria'. Tudi križanje 'Saikai 20' x 'Romantica' je imelo veliko stopnjo transgresivnih segregacij.

**Ključne besede:** soja; kombinacijsko žlahtnjenje; absolutna masa semen

1 Experimental Station on Soybean, Agricultural academy, Pavlikeni Bulgaria

2 Corresponding author, e-mail: gmvg@abv.bg

3 AgroBioInstitute, Agricultural Academy, Sofia, Bulgaria

## 1 INTRODUCTION

Legumes are widespread around the world, providing more than 69 % of the protein, and 30 % of the fats/oils needed for human diet. In the grain legumes - pea, soybean and common bean the protein content ranged from 20 % to 40 % depending on genotype and environment. Important is the fact that legumes require minimal amounts of soil fertilizers because they possess the ability to absorb nitrogen from the air through symbiotic interactions with nitrogen fixing bacteria. This ability affirm them as the world's plant protein sources, providing sustainable agriculture. Climate specificity of Bulgaria could be described as a heavy continental. The Bulgarian soybean varieties are created for the specific soil and climatic conditions of the country along with the developed agro-technologies. In Bulgaria the Experimental Soybean Station in Pavlikeny is the research organization where the Bulgarian soybean varieties are created and its main scientific activity is to develop varieties adapted to the country-specific agro-climatic environment and with high nutritional and technological quality of the grain. Technological qualities of the grain, including its specific absolute mass, play an important role in the purposeful use of soybean for food production. The mass of 100 seeds is an important component in seed yield, which is monitored in breeding productivity programs. The mass of 100 seeds possess complex genetic control. According to the study of Assefa et al. (2019), 14 quantitative trait loci (QTL) related to this trait were identified. Independent selection by absolute seed mass is considered easy and effective because of medium to high inheritance estimates (Cober et al., 1997; Krisnawati and Adie, 2015), but it is also important to associate this trait with productivity and adaptive potential. In their study (Mian et al., 1996) comment significant positive genotypic correlations between seed mass and seed yield per plant. On the other hand, the trait large seed is negatively correlated with the number of pods formed by the plant and the number of seeds per pod (Miladinović et al., 2011). Because these are the main criteria for selection in the direction of productivity, often high-yielding varieties are not large-seeded. It is important to note that seed mass is also strongly influenced by environmental factors. According to our research, genotypically defined trait large seed have a phenotypic realization when the reproductive phases of grain filling R7-R8 take place outside the period 1-15 August, when the drought period in Bulgaria is most frequent and most severe. In this regard, high and stable trait expression was observed in varieties of very early (MG 000-00) and medium to late

(MG I) maturity groups (Naydenova and Georgieva, 2019).

The nature of inheritance of the absolute mass of the seeds, established by the dialle crosses of the varieties grown in Bulgaria, is defined as additive (Kien, 1989). This is a reason transgressive selection to be considered as a potentially effective method to improve the trait. Transgressive variability in hybrid  $F_2$  and  $F_3$  families due to the recombination of additive alleles has been observed and further used in our selection program for all quantitative traits structuring seed yield in soybean (Aleksieva, 2001). According to Jambormias et al. (2015) the finding of multiple transgressive segregates at the  $F_3$  generation will enables an early selection to be carried out simultaneously for several multiple traits on self pollinated crops.

The purpose of this study is to identify positive transgressive forms by absolute mass of seeds, as well as to determine the potential of specific crosses and recombinant lines in the combinative breeding of high yielding large-seeded soybean varieties.

## 2 MATERIAL AND METHODS

### 2.1 CHARACTERISTICS OF PARENTAL GENOTYPES

Romantica variety is an ultra-early mature Ukrainian variety (MG00), with a vegetation period of about 90 days for the condition of Northern Bulgaria. In three years comparative variety trial, the value of the trait mass of 100 seeds for this genotype is 15.22 g (Naydenova and Georgieva, 2019). Unfavorable characteristics of the variety are low growth habitus and low positioning of the first pods.

Oria variety is an early mature Canadian variety (MG0) with a corresponding average mass of 100 seeds 14.14 g. Filling of the seeds in this genotype varies greatly in the years.

Felix variety is an early mature Romanian variety (MG0) with a mass of 100 seeds 13.27 g. It is characterized by a stable phenotypic expression of the trait, as well as high grain productivity.

Srebrina is a Bulgarian medium-early variety (MGI) with high adaptive and productive potential. The mass of 100 seeds for this genotype identified in the above mentioned comparative variety trail is 12.10 g, with trait values varying slightly over the years.

'Saikai 20' is a Japanese variety, middle-late maturity group (MGII), characterised with high-protein content. The genotype is small-seeded (mass of 100 seeds

– 11.15 g), with cracked pods under Bulgarian environmental conditions.

## 2.2 FIELD EXPERIMENTS

The crosses performed with these varieties were conducted in 2015 in the region of Pavlikeni (43°24'N, 25°32'E, 144 m), which is in the temperate continental climate zone with a well-established continental rainfall regime - with a maximum in May-June and a minimum in August-September. In 2017 the selection in  $F_2$  was made according to the growth habitus, by early maturity, height of positioning of the first pods, and not crackable pods. Generations  $F_3$  and  $F_4$  were tracked down in a breeding nursery compared to the parental genotypes over two consecutive years (2018-2019). The offspring of each elite genotype harvested in  $F_2$  generation was sown in one separate row 4 m in length, with a row spacing of 70 cm and 10 cm in row distance, 10 plants per linear meter.

## 2.3 PHENOTYPIC MEASUREMENTS AND DATA ANALYSIS

On the basis of visual assessment, 4 plants of each row were selected (10 % intensity of selection was applied), and the values of the trait mass of 100 seeds (g) and the yield of seeds per plant (g) were monitored for 20 plants of each family/studied cross. The mass of 100 seeds was evaluated in duplicate for each genotype. From the biometric measurements carried out, the mean values, limit values and variance by absolute mass of the seeds in  $F_3$  generation were determined. The positive transgressive forms of the studied trait were established in crosses, respectively families. Transgression is represented by the relative difference of the trait value of the individual hybrid combinations against that of the large seeded parent. On the base of the phenotypic data of the trait large seed (mass of seeds) in combination with the production of seed per plant, 46 lines were selected, and traced down in the fourth generation for stabilization of the trait large seed. The lines were sown in rows of 2 m in length, at 70 cm between rows and 5 cm in row distance, 20 plants per linear meter, in two randomized repeats. In the  $F_4$  generation, from each line, eight genotypes were selected, which were also subjected to biometric analysis.

Data from both generations ( $F_3$  and  $F_4$ ) were used to evaluate genotypic diversity and additive variance by absolute seed mass by family, respectively at specific crosses. Genotype diversity is represented by the inheritance coefficient  $H_{bs}^2$  calculated as the ratio of genotype to phe-

notype variance in  $F_3$  generation. The additive variance is represented by the inheritance coefficient  $h_{ns}^2$ , which is calculated by the parent-offspring covariance ( $F_4$ ) using the formula  $h_{ns}^2 = b/2r_{xy}$ , where  $b$  is the regression coefficient between the trait values in the parental genotypes ( $x$ ) and those in the offspring ( $y$ );  $r$  - is the correlation coefficient (Yankulov et al., 1993).

In order to evaluate the effectiveness of selection of the trait absolute seed mass in  $F_3$ , a rank analysis was performed. Spearman's rank correlation coefficient ( $r_s$ ) was used to estimate the relationship between seed mass of the elites selected in the  $F_3$  generation and that of their offspring in the  $F_4$  generation. The coefficient value is calculated using the formula given, where:  $d$  is the difference in genotype rank numbers by generation;  $n$  is the volume of the samples extract.

$$r_s = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}$$

Statistical and graphical data processing was done using Microsoft Excel 2010.

## 3 RESULTS AND DISCUSSION

According to the results (Table 1), the highest transgression rate by mass of 100 seeds is the hybrid combination in which both parental components are large-seeded - Romantica X Oria. For the family of this cross, a very high mean value of the trait was observed, as well as the highest degree of segregation with respect to its threshold values. Krisnawati and Adie (2015) had successfully obtained soybean large seed size lines established after crosses of large seeded parental genotypes (> 14 g/100 seeds). For the cross of the same maternal genotype with the Bulgarian variety Srebrina, high positive transgressions were also achieved, as well as a high range of phenotypic variability.

Using the Romantica variety as a pollen paternal component, lower mean values of the trait and significantly lower levels of positive transgressions were obtained. In these crosses, depending on the maternal component, contrasting results were also observed for variance of the trait. The highest hybrid variability was observed in the  $F_3$  population of the crossing breed where the Japanese variety Saikai 20 was the mother genotype. The results could be explained both by the large difference in trait values in parental components and that transgressive segregations are more likely in genetically distant genotypes (Stelkens et al., 2014). The 'Felix' x 'Romantica' cross family has the lowest variance in terms of the tracking trait.

**Table 1:** Mean values, limit values, variance and degree of positive transgression for families by trait mass of 100 seeds (g) in F<sub>3</sub> generation

	Hybrid combinations			
	'Romantica' x 'Srebrina'	'Romantica' x 'Oria'	'Felix' x 'Romantica'	'Saikai 20' x 'Romantica'
Mass of 100 seeds (m <sub>100</sub> ), g	19.3	21.9	18.5	17.6
Min	15.4	17.6	17.4	16.8
Max	23.2	26.4	19.5	20.4
σ <sup>2</sup> , g	5.3	10.9	0.8	13.0
Values of positive transgression	0.5 - 14.3%	1.4 - 36.1%	1.3 - 6.5%	0.2 - 10.8 %
Significance of genotypic variance in F <sub>3</sub> generation of family	$p < 0.01$	$p < 0.01$	$p = 0.10$	$p < 0.001$
H <sup>2</sup>	0.18	0.24	0.23	0.90

Investigation of heterogeneous F<sub>2</sub> and F<sub>3</sub> families, followed by testing only the best of the lines, is considered to be an effective method in early generation of soybean selection (St Martin and Gheraldi, 2002). In this case, in addition to the mean and limit values for the families studied, estimates of genotypic effect and genotypic diversity within them are also important. According to the variance analysis, the genotypic control in the variation of seed mass between the lines was significant ( $p < 0.01-0.001$ ) for three of four families/crosses studied - Table 1. This implies the efficiency of the selection by trait still in F<sub>3</sub> generation. On the other hand, according to the values of the inheritance coefficient, the genotypic diversity in F<sub>3</sub> generation of the crosses 'Romantica' x 'Srebrina', 'Romantica' x 'Oria', and 'Felix' x 'Romantica' is relatively low, where additional breeding criteria were used for the selection of elites (seed yield per plant, harvest index, growth habitus). The genetic differences in the studied generation of the 'Saikai 20' x 'Romantica' cross are significant ( $H^2 = 0.90$ ), which could be accepted as an indicator of both gene recombination leading to a significant increase in seed mass and the likely presence of spontaneous mutational variability of the trait.

It is important to consider the effect of environmental factors in determining the genotypic significance of the breeding trait in the individuals (Rosenzweig et al., 2016). The highest phenotypic variance established in families derived from parental components differing in maturity - 'Romantica' (MG00) x 'Srebrina' (MGI) and 'Saikai 20' (MGII) x 'Romantica' (MG00) - is possibly due to environmental impact. In the segregating hybrid populations, the individuals have different timing of grain-filling phenophases, which is a source of environmental variance by the traced trait (the differences in precipitation during the phenophases of grain filling R<sub>7</sub>-

R<sub>8</sub> in very early and middle-early offspring amounts to almost 50 mm). This requires the seed mass to be evaluated in relation to early maturity in these crosses. At the same time, high phenotypic variability by trait, coupled with a slightly segregation in maturity, is observed for the cross of early maturity 'Romantica' (MG00) x 'Oria' (MG0) varieties. In this case, some of the variability may be associated with a specific genotypic response to the larger feeding area in which the F<sub>3</sub> offspring were grown - row spacing 70 cm and 10 cm in row distance, 10 plants per linear meter. In our previous studies based on standard Bulgarian soybean varieties we found a difference in genotypic response of the trait seed mass related to sowing density (Georgiev et al., 2019). In this regard, an additional evaluation of the expression of the trait in crops with higher density is required, which was done in F<sub>4</sub> generation of crosses.

According to the results of the determining of phenotype by seed mass in combination with the seeds production per plant (Figures 1) higher frequency of offspring, significantly exceeding the parental forms is observed for the fourth hybrid generation of the cross 'Romantica' x 'Srebrina' - Fig. 1A, Fig. 2A. In relation to the results for the previous generation of this cross, the large seeded is connected to the maturity period, and the highest seed mass was observed for the offspring of the late mature F<sub>3</sub> genotypes. For genotype R5/13/2 an extremely high mass value of 100 seeds was reported - 28.9 g.

In F<sub>4</sub>, the offspring of the cross 'Romantica' x 'Oria', under conditions of genotypic and environmental competition, were observed only two elites, superior in terms of both productivity and seed mass of the parent components (Fig. 1B, Fig. 2B). It is important to note that there are also a large number of offsprings with very high matter of the breeding trait (>20 g), which exhibit higher in-

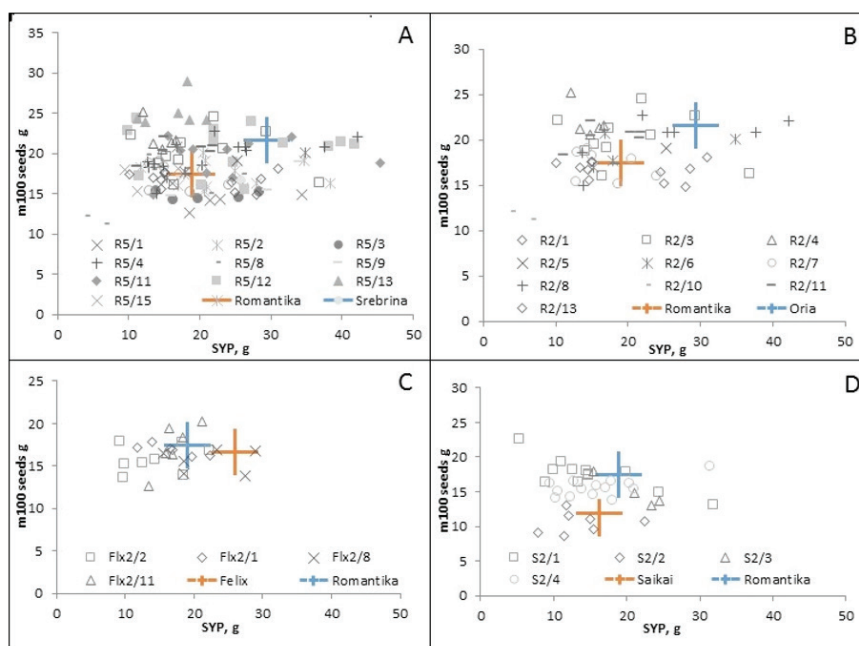


Figure 1: Scattering of seed yield per plant (SYP, g) by mass of 100 seeds ( $m_{100}$ , g) in F4 generation of crosses (A-D)

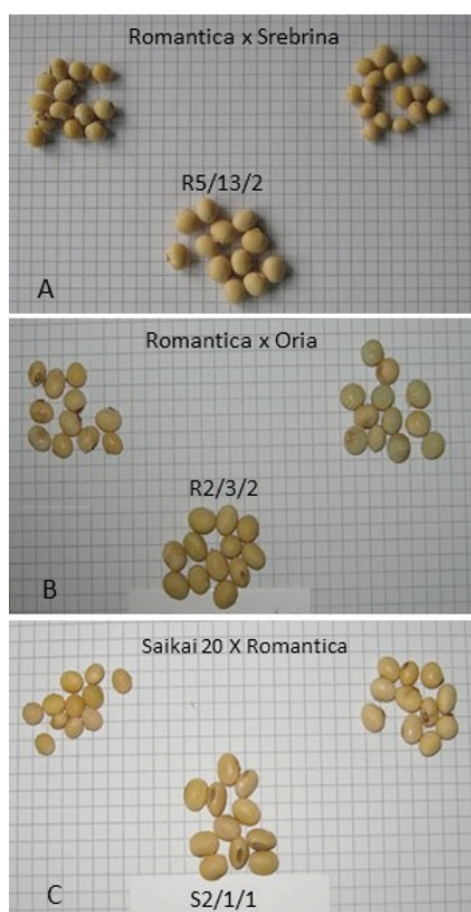


Figure 2: Representative elite lines from F4 generation (A-C)



**Table 2:** Correlation coefficient between mass of 100 seeds ( $m_{100}$ ), seed yield per plant (SYP) and number of seeds per plant (NSP) in  $F_4$  generation, rank correlation coefficient by trait  $m_{100}$  values in  $F_3$  and  $F_4$  generations ( $r_s$ ), inheritance coefficient of  $m_{100}$  ( $h_{ns}^2$ )

Hybrid combinations	'Romantica' x 'Srebrina'	'Romantica' x 'Oria'	'Felix' x 'Romantica'	'Saikai 20' x 'Romantica'
$r_{m100/SYP}$	0.13	0.07	0.12	-0.22
$r_{m100/NSP}$	-0.28	-0.10	-0.15	-0.62
$r_s$	0.60	0.69	0.54	0.80
$h_{ns}^2$	0.47	0.56	0.39	0.54

dividual productivity than that of the mother variety. This allows a high intensity of selection by the mass of seeds in the family.

Due to the low hybrid variability presented above, the number of 'Felix' x 'Romantica' progeny monitored in  $F_4$  is small. According to the results, the cross has no potential for the desired combination of high seed mass with high plant seed yield - Fig. 1 C.

In a large number of  $F_4$  offsprings from the 'Saikai 20' x 'Romantica' cross, a significant increase in the selected trait values was observed against the maternal genotype. In the offspring of one of the selected line in the previous generation, genotypes with very high seed mass values (> 20 g) and with high yield of seeds per plant (> 30 g) were observed (Fig.1 D, Fig.2 C). Genotypes have been found to exceed the Romantica variety for both traits, which is the parent component with higher meaning of these traits.

The efficiency of identifying offspring with high genotypic meaning by mass of 100 seeds in  $F_3$  estimated by the rank correlation coefficient ( $r_s$ ) is very high for the 'Romantica' x 'Oria' and 'Saikai 20' x 'Romantica' crosses - Table 2. Also relevant are the results of the assessment of additive inheritance by the trait in the commented crosses - the inheritance coefficient  $h_{ns}^2$  possess corresponding values of 0.56 and 0.54, the additive variance is about half of the observed phenotypic variance.

According to the values of the correlation coefficient ( $r_{m100/SYP}$ ), the mass of the seeds was not related to the yield of seeds per plant (Table 2). A significant compensating dependency between seed mass and number of seeds per plant ( $r_{m100/NSP} = -0.62$ ) was observed for recombinant lines from the 'Saikai 20' x 'Romantica' cross. According to Fujii et al. (2018) some genetic factors for large seeds reduce the number of pods and fertility by decreasing either the number of pods per plant or the number of ovules per pod. Definitely further studies are needed to find out whether the regulation of these components is due to a common genetic factor or closely related ones.

#### 4 CONCLUSION

According to the results presented in this study,

transgressive selection can be successfully used to reach the goal large seeds in soybean. The efficiency of selection of transgressive forms in  $F_3$  generation is high. The valuable combinative variability in seed mass obtained in all the crosses studied, identifies the Romantica variety as an important genetic source in the selection for the trait large seed. The genetic potential to combine a high specific mass of seeds with a high yield of seeds per plant has been established for the 'Romantica' cross with the Bulgarian standard variety Srebrina.

Recombinant lines suitable for intensive selection for the trait large seed were obtained from the 'Romantica' x 'Oria' combination. The 'Saikai 20' x 'Romantica' cross possess a very high degree of transgressive segregations. The hybrid material derived from it could be successfully used to study the genetic factors affecting both traits - the number of seeds per plant and their specific mass.

#### 5 ACKNOWLEDGEMENT

This study was supported by the Bulgarian Ministry of Education and Science under National Research Programme „Healthy Food for Strong Bio-Economy and Quality of Life“ approved by DCM №577/17.08.2018.

#### 6 REFERENCES

- Aleksieva, A. (2001). Study of transgressive segregations on some quantitative traits of soy. *Scientific Papers of AU Plovdiv, XLVI* (3), 129-134.
- Assefa, T., Otyama, P. I., Brown, A. V., Kalberer, S. R., Kulkarni, R. S., & Cannon, S. B. (2019). Genome-wide associations and epistatic interactions for internode number, plant height, seed mass and seed yield in soybean. *BMC Genomics*, 20(1), 527. <https://doi.org/10.1186/s12864-019-5907-7>
- Cober, E. R., Voldeng, H. D., & Fregeau-Reid, J. A. (1997). Heritability of seed shape and seed size in soybean. *Crop Science*, 37(6), 1767-1769. <https://doi.org/10.2135/cropsci1997.0011183X003700060017x>
- Fujii, K., Sayama, T., Takagi, K., Kosuge, K., Okano, K., Kaga, A., & Ishimoto, M. (2018). Identification and dissection of single seed mass QTLs by analysis of seed yield components

- in soybean. *Breeding science*, 68(2), 177-187. <https://doi.org/10.1270/jsbbs.17098>
- Georgiev, G., Naydenova, G., Todorova, R. (2019). Effects of the sowing time and row spacing on the structural elements of yield for two soybean varieties. *Field crop studies. XII* (3).
- Hu, Z., Zhang, H., Kan, G., Ma, D., Zhang, D., Shi, G., Hong, D., Zhang, G., Yu, D. (2013). Determination of the genetic architecture of seed size and shape via linkage and association analysis in soybean (*Glycine max* (L.) Merr.). *Genetica*, 141, 247-254. <https://doi.org/10.1007/s10709-013-9723-8>
- Jambormias, E., Sutjahjo, S. H., Mattjik, A. A., Wahyu, Y., Wirnas, D., Siregar, A., Patty J. R., Laisina J. K., Madubun E. L. & Ririhena, R. E. (2015). Transgressive segregation analysis of multiple traits in mungbean (*Vigna radiata* (L.) Wilczek). *SABRAO Journal of Breeding & Genetics*, 47(2), 201-213.
- Kien, T. D. (1989). *Variability, inheritance and correlations between important agronomically characteristics of soybeans*. PhD thesis, Institute of Genetics, BAS.
- Krisnawati, A., & Adie, M. (2015). Selection of soybean genotypes by seed size and its prospects for industrial raw material in Indonesia. *Procedia Food Science*, 3, 355-363. <https://doi.org/10.1016/j.profoo.2015.01.039>
- Mian, M., Bailey, M., Tamulonis, J., Shipe, E., Carter, T., Parrott, W., Boerma, H. (1996). Molecular markers associated with seed mass in two soybean populations. *Theoretical and Applied Genetics*, 93(7), 1011-1016. <https://doi.org/10.1007/BF00230118>
- Miladinović, J., Burton, J. W., Tubić, S. B., Miladinović, D., Djordjević, V., & Djukić, V. (2011). Soybean breeding: comparison of the efficiency of different selection methods. *Turkish Journal of Agriculture and Forestry*, 35(5), 469-480.
- Naydenova, G., & Georgieva, N. (2019). Study on seed yield components depending on the duration of vegetation period in soybean. *Bulgarian Journal of Agricultural Science*, 25(1), 49-54.
- Rosenzweig V.E., Goloenko D.V., Davydenko O.G., (2016). Selection in heterogeneous soybean populations: Identification of valuable genotypes Oil crops. *Scientific and Technical Bulletin of the All-Russian Research Institute of Oil crops*, 2(166), 26-33.
- St Martin, S. K., & Geraldi, I. O. (2002). Comparison of three procedures for early generation testing of soybean. *Crop science*, 42(3), 705-709. <https://doi.org/10.2135/cropsci2002.7050>
- Stelkens, R. B., Brockhurst, M. A., Hurst, G. D. D., Miller, E. L., & Greig, D. (2014). The effect of hybrid transgression on environmental tolerance in experimental yeast crosses. *Journal of evolutionary biology*, 27(11), 2507-2519. <https://doi.org/10.1111/jeb.12494>
- Wallace, D. H., Baudoin, J. P., Beaver, J., Coyne, D. P., Halseth, D. E., Masaya, P. N., Zobel, R. W. (1993). Improving efficiency of breeding for higher crop yield. *Theoretical and Applied Genetics*, 86(1), 27-40. <https://doi.org/10.1007/BF00223805>
- Yankulov, M., Daskalov, S., Tomov, N., Atanasov, A., Vitanov, M., Roseva, A., Lidanski, T., Georgiev, H., Achkova Z. (1993). Principles of modern selection. *Zemizdat, Sofia*, 274 pages.