Bulgarian Journal of Agricultural Science, 30 (No 6) 2024, 1050–1058

Comparison of yield components and detection of seed size associate locus SW9-1 in elite soybean breeding lines

Katerina Stefanova[1], Mariana Radkova[1*], Galina Naydenova[2] and Anelia Iantcheva[1]

[1]*Agricultural Academy, Agrobioinstitute, 1164 Sofia, Bulgaria*
[2]*Agricultural Academy, Research Institute of Mountain Stockbreeding and Agriculture, 5600 Troyan, Bulgaria*
[*]*Corresponding author:* marianaradkova@yahoo.com

Abstract

Stefanova, K., Radkova, M., Naydenova, G. & Iantcheva, A. (2024). Comparison of yield components and detection of seed size associate locus SW9-1 in elite soybean breeding lines. *Bulg. J. Agric. Sci., 30*(6), 1050–1058

Soybean is one of the main sources of protein for humans and livestock. Soybean seeds are also very rich of oils. Seed size is important parameter in all agricultural crops, and as the main component of yield is the subject of research in breeding programs. In Bulgaria, soybean crop is grown without irrigation, which requires the cultivation of early varieties to allow the flowering and the filling of the seeds to take place before the extreme high temperatures and scarce rainfall during the hottest summer months July and August. For this reason the selection is conducted in both directions – large-seeds and early maturity. Modern selection methods such as marker-assisted selection shorten the time for developing of breeding materials.

In the present study, genotypic diversity related with the seed productivity and the factors determining it, were investigated in F5-F6 soybean breeding lines. The lines were realized after four different crosses including very early cultivar – Romantica; cultivar with high protein content – Saikai and cultivars Srebrina and Galina forming big seeds. The possibilities of the SW9-1 locus carrying SNP related to seed size in soybean to serve as a marker in a selection process was also investigated.

Keywords: Glycine max (L.) Merrill; soybean breeding; seed size; SNP

Introduction

The main goals of the breeding programs are to improve the crops traits related to yield, nutrition quality and bet ter adaptation to environmental changes. Generally the in creased yield of soybean is a result of growing land areas use, and a small part is related with the selection improve ment of the soybean germplasm. This is the main difference between soybean and cereal crops.

Shifting the critical reproductive phenophases of soybeans to periods of better moisture security through agrotechnolog ical approaches (cultivation of early maturing varieties, very early or late sowing, etc.) is a major possibility for stabiliz-ing crop yield under non-irrigated cultivation in Bulgaria. Therefore, when selecting genetic sources for combinative breeding, seed productivity and its structuring components are studied in relation to the maturity group of the genotypes in-cluded in the crosses. In the breeding program related to large seed, genotypes with high expression of the trait from the ear-ly or mid-late maturity group were used as parental compo-nents. The average values, but also the ecological dispersion of the indicators like number of fruiting nodes and number

of branches are higher in the genotypes with a longer growing season. It was found that some early-ripening genotypes also demonstrated high and ecologically stable expression of these parameters (Naydenova & Georgieva, 2019). The established genotypes under the conditions of Southern Europe with Bulgarian, Serbian, Ukrainian and French origin, for seed production, show stable appearance of individual productivity, assessed by the number and weight of seeds per plant, as well as the parameters harvest index and absolute seed mass (Naydenova & Georgieva, 2019; Radkova & Naydenova, 2022). These characteristics were demonstrated for early and mid-early varieties. Therefore, such genotypes have been used as paternal components in our improvement work aimed for selection of high-yielding, large-seeded and high-protein content soybean varieties from the early maturity group.

Seed size is one of the main agronomic characteristics for each crop, because of the direct relation with the yield. Soybean seeds weight is presented as a 100-seed weight (100-SW) and is affected by seed size. The soybean yield and 100-SW are affected by multiple quantitative trait loci (QTL) or genes (Mansur et al., 1993; Wang et al., 2015) which make it difficult to improve soybean yield by traditional breeding methods (Hao et al., 2012). This is the reason to support the breeding programs by application of molecular markers (marker assisted selection).

The molecular markers are fast and accurate approach for identification of new sources of genetic diversity as well as a suitable tool for discovery of the genetic factors controlling the quantitative traits (Mian et al., 1996; Sun et al., 2012).

Soybean possesses low genetic diversity base on morphological and RFLP (Restriction Fragment Length Polymorphism) data (Keim et al., 1992; Shoemaker et al., 1992, cited by Hudcovicova & Kraic, 2003). This is the reason for testing different molecular markers such as RAPDs (random-amplified polymorphic DNA), SSRs (simple sequence repeats) and others in order to study genetic diversity especially in these genotypes, which are morphologically identical. In the publications of Song et al. (2010), Hipparagi et al. (2017) and others authors, the application of SSR markers in soybean was described. According to Cox et al. (1985) and Messmer et al. (1993), cited by Hudcovicová & Kraic (2003), genotyping of cultivars with similar agronomic traits, but genetically different could produce highly variable progenies by heterosis effect. That way, the criteria for selection of suitable parental genotypes has to be based not only to the agronomic value, but also according to their genetic diversity.

The locus SW9-1 located on chromosome 9 from the soybean genome was published as a new locus related significantly to the seed size (Li et al., 2019). The authors found that this locus has high phenotypic effect over the trait 100-SW, and often is related to the genotypes with large seeds. They also demonstrated that the increase of the seed weight is accompanied with the presence of the SW9-1T allele. This locus has been subjected to artificial selection during the early stages of soybean breeding, especially the utilization of SW9-1T in the cultivar Edamame for big seeds. The locus was detected by application of genome – wide association study and genotyping of single sequence polymorphisms (SNP) (Li et al., 2019).

The following paper presents the genotypic diversity in terms of seed productivity and the factors determining it in selected F5-F6 lines with a high degree of intensity from the following crosses: ♀Romantica X ♂Oria (R2), ♀Romantica X ♂Srebrina (R5), ♀Saikai X ♂Romantica (S2), ♀Saikai X ♂Galina (S3). Based on the results, five elite F7 lines – R2, R5, S2, S3, S32 were selected and subjected to genetic analysis, with the aim to select the best of them as new cultivars. Additionally, two lines out from the selected group were analyzed in order to detected relation between large seed trait and the presence of SNP in the locus SW9-1.

Material and Methods

Phenotypic studies

In two consecutive years (2020-2021) F5-F6 early maturing (MG0) recombinant lines from the varietal crosses ♀Romantica X ♂Oria (R2), ♀Romantica X ♂Srebrina (R5), ♀Saikai X ♂Romantica (S2) and ♀Saikai X ♂Galina (S3) as well as the standard variety Avigea were sown in rows 4m long with an inter-row distance of 70 cm and an intra-row distance of 5 cm, in two randomized replications. The following yield components were traced: plant height (cm); number of branches; number of pods per plant; number of seeds per plant; seed weight per plant (g); absolute mass of seeds (g). Biometric measurements were determined in 10 plants of each line. The phenotypic diversity for the studied traits in each generation was characterized by mean value (x) and standard deviation (SD). Data were processed by one-factor analysis of variance. Significant difference between mean value was determined by LSD test with a significance level of 0.05. Genotypic diversity was estimated by the heritability coefficient Hbs^2, calculated as the ratio of genotypic to phenotypic variance: $H^2bs = s^2_G/s^2_{Ph}$. From F5 generation were analyzed only these lines from which F6 recombinant early-maturing lines were produced. These lines were characterized as elite according to a complex selection evaluation compared to the standard variety Avigea.

The main tendency and the range of the frequency distribution according to the selection criteria in the F6 populations of each cross are presented graphically.

Molecular analysis

Genomic DNA was extracted from ten individual 4 weeks old plants, from the soybean cultivars: Avigea, Srebrina, Romantica, Saikai and Galina and lines S2/4/3/8 and R5/11/11/7. The collected leaf tissue was grounded with Tissue LyserII (Qiagen). The protocol of Plant & Fungi DNA purification kit (EURx) was followed. The DNA concentration was measured spectrophotometrically with NanoDrop 2000, (Thermo) and diluted up to 50 ng/µl.

Amplification conditions were according the protocol of Li et al. (2019). Briefly 1µl DNA was mixed with 12.5 µl 2xHS My Taq mix (Bioline), 1 µl dCAPS *XbaI* F primer, 1 µl dCAPS *XbaI* R primer and ultrapure water to 25 µl reaction volume. PCR reaction was following the steps: denaturation at 95°C for 3 min; 35 cycles of denaturation at 95°C for 15 s, annealing at 55°C for 45 s (-0.2°C at each cycle) and extension at 72°C for 7 min. The final extension step was conducted at 72°C for 3 min. Five microliters from PCR reaction was monitored on agarose gel. The rest part of the reaction was digested with *XbaI* restriction enzyme. After enzyme digestion the reactions were again monitored on the gel. The expected correct fragments with length of 200 bp were purified from the gel following the protocol of GeneJET PCR Extraction Kit (Therno Scientific) and used for cloning. Cloning procedure was according to the protocol of CloneJET PCR Cloning Kit (Therno Scientific). Plasmids DNA was extracted from positive recombinant *E. coli* clones and sequenced by Macrogene company (https://dna.macrogen.com/main.do#). The obtained sequences were aligned with the program vector NTI 11.5.

Results and Discussion

Phenotypic analysis

The genotypic variance of yield-determining traits was significant in both monitored generations of the selected early-maturing recombinant soybean lines. An exception was observed for the trait weight of seeds obtained from a plant in the F6 generation (Table 1). According to the values of the inheritance coefficient, the greatest genotypic diversity was observed for the traits plant height and absolute seed mass.

The parameter plant height is mostly often negatively related to early maturity (MG0), used by us as the main selection criteria in previous generations. This can explain the results obtained for the F5 generation, where all lines are shortened with significant differences to the standard variety Avigea, which belong to the medium-early maturity group (MG0-I) (Table 2).

In the F6 generation, the indicated negative dependence was overcome for the line R5/12/12/1, for which a significant superiority ($P<0.05$) in plant height compared to Avigea was observed (Table 3). The combination of greater plant height, respectively greater number of fruit nodes with early maturity is considered as an effective approach to increase the productive potential in soybean (Borowska & Prusiński, 2021). In this regard, the observed high genotypic diversity in terms of plant height in the group of lines with maternal genotype Romantica, which is very early-ripening variety (MG0) and possesses a short – growing habitus.

The amount of branches formed per plant is also an important component of grain yield in soybean, with a clear tendency for later maturing genotypes to form more branches (Naydenova & Georgieva, 2019; Liu et al., 2020). The high productive potential of the standard variety Avigea is connected in the first place with the large number of branches of the first order (Aleksieva, 2015). In the current study, the largest number of branches was reported for cultivar Avigea, and two lines (F5 generation) from the combinations with the maternal component – variety Romantica. These lines fall into homogeneous group of the standard variety Avigea. In F6 generation, large number of branches was observed for three more lines with maternal genotype Saikai, which belong to the late maturity group (MGII). The heritability coefficient has values of 21% and 29% for F5 and F6 genera-

Table 1. Components of variance and heritability coefficient H_{bs}^2 regarding the structural elements of yield in soybean lines

Genera-tions	Variance of heritability coefficient	Plant height, cm	Number of branches	Number of pods/plant	Number of seeds/plant	Seed mass/ plant, g	100 seeds mass, g
F5	S^2g	192.90***	0.27**	na	806.77***	61.65***	9.42***
	S^2ph	235.79	1.26	na	2966.14	116.13	11.83
	H_{bs}^2	0.82	0.21	na	0.27	0.53	0.79
F6	S^2g	176.04***	0.447**	77.29*	352.21**	2.10^{ns}	893.22***
	S^2ph	202.3	1.55	408.19	1263.60	17.83	1192.80
	H_{bs}^2	0.87	0.29	0.19	0.28	0.12	0.75

Table 2. Monitored parameters of selected lines from crosses in comparison with standard Avigea – F5 generation

F5	St	♀Romantica X ♂Oria		♀Romantica X ♂Srebrina			♀Saikai X ♂Romantica	♀Saikai X ♂Galina		
	Avigeia	R2/8/7	R2/8/1	R5/11/11	R5/1/1	R5/12/12	S2/4/3	S3/2/1	S3/1/5	S3/1/6
PH, cm Mean	94.3 a	73.4 c	58. 2e	63.8 de	72.2 c	86.6 b	58.6 e	70.3 cd	61.6 e	68.5 cd
SD	8.1	8.5	5.5	4.6	7.6	7.0	5.9	5.7	6.9	5.0
NB Mean	4.5 a	4.2 a	3.1 bc	2.8 c	3.2 bc	3.9 ab	2.5 c	2.8 c	2.7 c	2.8 c
SD	1.8	1.3	1.0	1.3	0.7	0.8	0.7	1.0	0.6	1.2
NS Mean	94.0 cde	128.6 bc	73.5 de	189.0 a	106.0 bcd	146.5 ab	66.3 e	102.3 bcd	116.2 bc	130.5 bc
SD	57.0	44.3	23.2	62.3	57.7	44.2	36.7	41.3	55.5	35.6
SW, g Mean	11.1 c	24.3 b	9.9 c	38.5 a	13.4 c	28.5 b	8.9 c	14.8 c	12.0 c	14.8 c
SD	6.0	7.3	3.1	16.6	8.0	9.6	6.6	6.5	6.8	3.8
HI, % Mean	44.4 abc	43.7 abcd	40.2 bcde	48.9 a	39.4 cde	44.7 ab	38.3 de	39.8 bcde	40.3 e	41.3 bcde
SD	2.9	3.6	1.9	4.3	3.3	3.8	3.4	4.9	4.9	3.0
m 100, g Mean	11.9 d	19.1 a	13.6 bc	19.9 a	12.6 cd	19.3 a	12.7 cd	14.2 b	10.2 e	11.4 de
SD	2.0	2.0	0.9	2.9	1.6	1.2	1.7	1.9	1.1	0.9

Legend: PH – plant height; NB – number of branches; NS – number of seeds; SW – seeds weight; HI – harvest index; m100 – mass of 100 seeds; SD – standard deviation

tions (Table 1), respectively, which is indicative of relatively low genotypic diversity in terms of number of branches per plant in the studied group of lines.

The number of formed pods and seeds per plant are difficult to be influenced by breeding, because of the high importance of the environmental factor on their phenotypic expression. Despite the low scores for genotypic diversity (Hbs^2 from 0.19 to 0.28), in the F5 generation two lines – R5/11/11 and R5/12/12 have significantly higher values of these indicators compared to the standard variety. In the F6 generation, all lines from the crosses of Romantica variety as the maternal genotype fall into a homogeneous group with the standard Avigea. A reliable superiority in comparison with the standard was observed for the lines S3/1/5/21 and S2/4/3/8 originating from the combinations with the Saikai variety as a maternal genotype. Since the evaluation of the two generations was carried out in two years, respectively under different climatic conditions, the divergent appearances of the individual lines suggest the presence of non-additivity – an effect of the interaction genotype-environment on the values of the indicators number of pods and seeds per plant. In this regard, the stability of the appearance of the indicated lines should be evaluated.

The individual productivity which is assessed by the weight of seeds obtained from a plant is also related to the results presented so far. In the F5 generation, the genotypic diversity is high – Hbs^2 is 0.53 and the genotypic differences, highly reliable ($P<0.001$). Three lines from the crosses with the Romantica as maternal genotype have significantly higher individual productivity compared to the standard – R5/11/11, R5/12/12 and R2/8/7. Contrasting results were observed in F6 – the genotypic variance was unreliable ($P=0.09$) and the genotypic diversity score was very low – $Hbs^2 = 0.12$. In the multiple comparison, only line R5/11/11/7 provenly exceed Avigea variety in terms of seed weight per plant, same as in the previous generation. All other genotypes fell into a homogeneous group in terms of individual productivity in the range 5.1-12.0 g. The significant superiority over the standard variety of lines S3/1/5/21 and S2/4/3/8 in the number of pods and seeds formed per plant does not give significant advantage in individual productivity, which in this case is related to small-seeded /lower absolute seed mass.

Absolute seed mass is an indicator with strong genetic determination. In both monitored generations, a very high genotypic diversity of the mass per 100 seed was observed between the selected recombinant lines. In F5, genotypic differences were more pronounced, being significant even between lines originating from the same combination. In three lines – R5/11/11, R5/12/12 and R2/8/7 very high values of the indicator mass per 100 seed >19.1 g were recorded. In F6, a significant excess compared to the standard variety was again observed in the lines from the crosses with the mater-

Table 3. Parameters monitored of selected lines from crosses in comparison with standart Avigeia – F6 generation

F6	St	♀Romantica X Oria		♀Romantica X ♂Srebrina					♀Saikai X ♂ Roman-tika	♀Saikai X ♂Galina				
	Avigea	R2/8/7/1	R2/8/2/1	R5/1/1/9	R5/11/11/2	R5/11/11/6	R5/11/11/7	R5/12/12/1	S2/4/3/8	S3/2/1/2	S3/1/5/10	S3/1/5/12	S3/1/5/21	S3/1/6/11
PH, cm Mean	70.6 bcd	74.6 b	52.7 e	71.8 bc	71.2 bcd	67.2 cd	75.2 b	105 a	67.3 cd	65.2[d]	56.6 e	57.0 e	50.6 e	54.2 e
SD	3.4	6.2	5.5	3.2	2.6	8.1	5.7	8.6	7.3	2.6	2.6	2.5	1.1	4.7
NB Mean	4.6 a	3.2 bc	3.0 bcd	2.8 bcd	1.6 d	2.2 cd	2.8 bcd	3.2 bc	3.8 ab	3.2 bc	3.0 bc	3.8 ab	4.6 a	3.2 bc
SD	0.9	1.1	1.0	1.1	0.9	1.1	1.3	0.8	1.0	0.8	0.7	1.8	1.1	0.4
NP Mean	39.2 d	32.8 d	56.0 abcd	41.4 cd	43.4 bcd	40.8 cd	48.0 bcd	37.4 d	62.8 abc	51.2[bcd]	53.0 abcd	54.4 abcd	74.4 a	65.6 ab
SD	9.2	4.7	6.1	17.1	8.4	13.6	28.8	13.1	30.0	16.8	10.1	21.8	24.3	17.3
NS Mean	70.0 de	60.4 de	93.3 abcd	48.2 e	67.8 de	73.2 cde	81.4 bcde	63.2 de	126.2[a]	88.4[bcd]	87.0[bcd]	99.3 abcd	117.6[ab]	108.8[abc]
SD	12.4	11.3	12.9	16.8	13.1	26.5	45.5	23.0	61.6	22.2	15.8	18.6	37.6	29.7
SW, g Mean	6.4 bc	7.8 bc	12.0 ab	5.1 c	9.5 abc	10.8 ab	13.5 a	9.0 abc	11.0 ab	9.0 abc	6.6 bc	8.6 abc	9.7 abc	8.7 abc
SD	2.3	1.9	2.4	1.9	2.8	3.9	9.3	1.7	7.0	2.9	1.7	0.9	3.9	2.0
m 100, g Mean	8.9 ef	13.6 abc	13.0 bc	11.5 cd	15.0 ab	14.8 ab	15.8 a	15.0 ab	8.4 ef	9.9 de	7.5 f	8.8 ef	8.0 ef	8.1 ef
SD	15.0	14.5	31.0	17.3	17.2	12.0	19.9	33.0	15.3	8.6	11.7	12.2	12.3	10.0

Legend: PH – plant height; NB – number of branches; NP – number of pods; NS – number of seeds; SW – seeds weight; m100 – mass of 100 seeds; SD – standard deviation

nal large-seeded genotype Romantica. When a small-seeded variety (Saikai) was used as the maternal genotype, the genotypic variance for the trait was insignificant, and the lines, together with the standard variety Avigea, fell into a homogeneous group characterized by small-seeds.

In summary, the advantages of the elite recombinant lines analyzed in two generations compared to the standard variety are in terms of number of pods and seeds per plant (for the lines from the ♀Saikai cross), as well as in individual plant productivity and specific seed mass (in lines from the ♀Romantica X ♂Srebrina cross).

The importance of the maternal genotype was determined in the characterization of the individual recombinant lines. Of interest is the variance in the observed traits depending on the paternal genotype. Figure 1 (A-E), presents the main tendency by mean and median, as well as the range of the frequency distribution by selection criteria in the F6 populations of each cross. The comparison between the two populations with the maternal genotype Romantica shows that the paternal genotype affects the main tendency and dispersion in the number of branches per plant, which are higher in the combination of Romantica with Oria. In this cross,

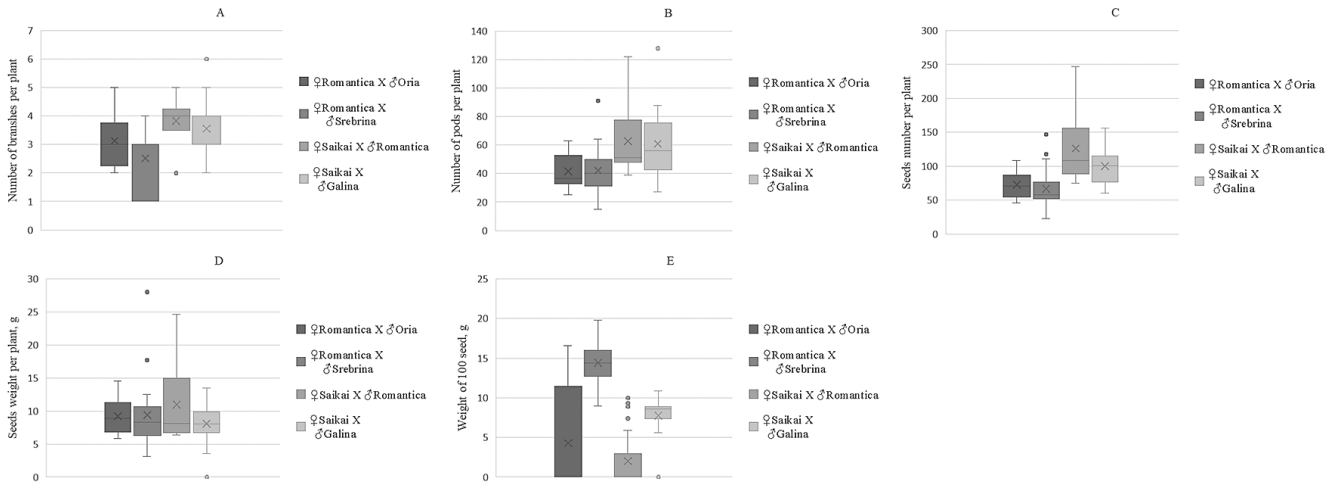

Fig. 1. The main tendency and the range of the frequency distribution according to the number of branches per plant (A), pods per plant (B), seeds number per plant (C), seeds weight per plant (D) and weight of 100 seed (E) in the F6 populations of each cross

both parental genotypes belong to early maturity group and form large seeds.

Regarding the number of pods and seeds per plant, absolute seed mass, as well as individual productivity, the two populations are relatively uniform in average values and interquartile range, but when using the Srebrina variety as paternal genotype, significantly higher values of recombination variability were achieved. Lines were observed and selected accordingly, significantly standing out with their high values of the commented indicators.

In the crosses with the Saikai variety as a maternal genotype, the differences between populations depending on the paternal genotype are more pronounced. With higher values of indicators of main tendency and dispersion, correspondingly higher selection value in number of seeds, absolute seed mass and seed weight per plant is the F6 population of the cross Saikai x Romantica were achieved. This is the combination with the highest genetic distance – the Saikai variety is late, small-seeded, origin of Asia, and the Romantica variety is very early-ripening, large-seeded, origin of Eastern Europe. In the second combination of the Saikai variety with the Galina variety as a parent, significant variability was observed in individual genotypes in terms of the number of branches and pods per plant and absolute seed mass.

Molecular studies

In the following study two selected lines – S2/4/3/8 and R5/11/11/7 were analyzed for the structure of *XbaI* site located in the region of locus SW9-1 in comparison with parental genotypes and the standard cultivar Avigea. Both lines S2/4/3/8 and R5/11/11/7 are realized from the cross between more genetically distant parental genotypes. Line S2/4/3/8 originated from cross between Saikai as mother parent – late genotype characterized with small seeds and Romantica – very early cultivar (MG0) with large seeds. Line R5/11/11/7 is realized from the cross between Romantica cultivar as mother plant and Srebrina – Bulgarian middle early (MGI) cultivar, as a father plant. The results using CAPS marker for SNP detection after enzymatic cleavage by *XbaI* (Li et al., 2019) show the presence of the mutated allele from C to T, resulting in no restitution, in one plant out of 10 tested from the cultivars Srebrina and Avigea (Figures 2 and 3). Additionally one heterozygous plant by this allele from cultivar Srebrina was detected – lane 6, (Figure 2). The variety Romantica with seed weight of 23.5 g, the mutant allele T

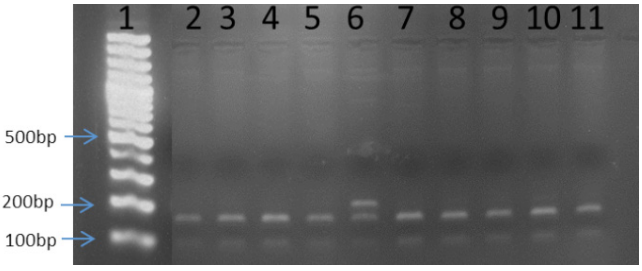

Fig. 2. PCR products from cultivar Srebrina after *XbaI* enzyme digestion. 2.5% agarose gel electrophoresis

Legend: 1. DNA leader; from to 2 to 11 – individual plants from cultivar Srebrina. On position 6 is presented the pattern of the heterozygous plant

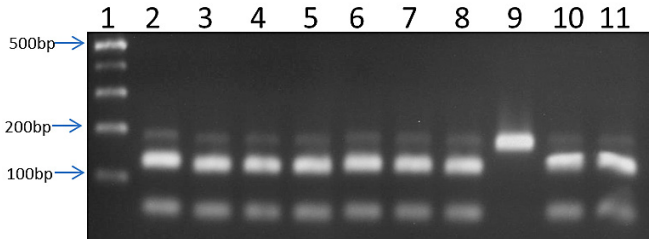

Fig. 3. PCR products from cultivar Avigea after *XbaI* enzyme digestion. 2.5% agarose gel electrophoresis
Legend: 1. DNA leader; from to 2 to 11 – individual plants from cultivar Avigea. On position 9 is presented the pattern of undigested plant

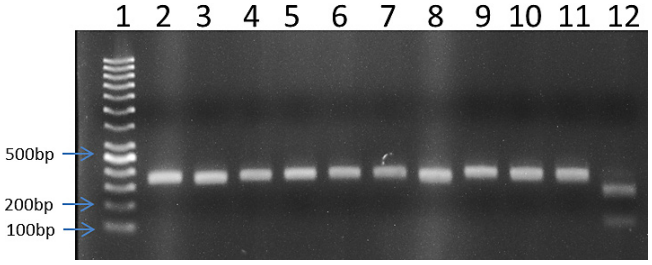

Fig. 4. PCR products from cultivar Romantica after *XbaI* enzyme digestion. 2.5% agarose gel electrophoresis
Legend: 1. DNA leader; from to 2 to 11 – individual plants from cv. Romantica; 12. Positive control – digested R5 plant

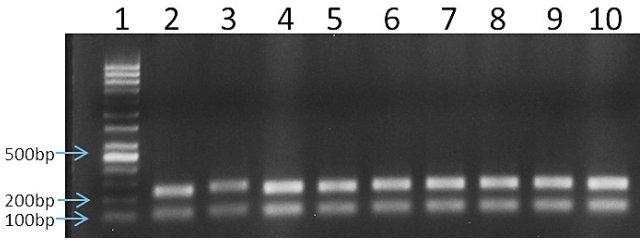

Fig. 5. PCR products from line R5 11/11/7 after *XbaI* enzyme digestion. 2.5% agarose gel electrophoresis
Legend: 1. DNA leader; from to 2 to 10 – individual plants from line R5

was detected in all (ten) individual plants tested related with no digestion profile (Figure 4). It was interesting to find that the genotype Saikai, which forms small seeds (13.1 g) the mutant T allele was also detected in all (nine) individual plants tested (data not shown). Same digestion analysis was conducted for ten individual plants from the lines S2/4/3/8 and R5/11/11/7 (Figure 5). The pattern of both lines demonstrated absence of the mutant allele and all individual plants were able to digest with *XbaI*.

In order to confirm our results the obtained PCR products (one plant per genotype) were purified, cloned and sequenced. From Srebrina cultivar both alleles were cloned from the heterozygous plant. The alignment of all sequences is presented on Figure 6. The results confirm the presence of the mutated nucleotide from C to T in the *Xba I* site in the following genotypes: Avigea, Srebrina, Romantica and Saikai. In the lines

```
                1                                                          73
       R5     (1) TGTTTGAATGGTGGGGTTAAGGAAAGAAGAAGAAATTATTAGGTTTTTTTTGTGAGCAAGAAGAAATTATTAGG
   Galina     (1) TGTTTGAATGGTGGGGTTAAGGAAAGAAGAAGAAATTATTAGGTTTTTTTTGTGAGCAAGAAGAAATTATTAGG
       S2     (1) TGTTTGAATGGTGGGGTTAAGGAAAGAAGAAGAAATTATTAGGTTTTTTTTGTGAGCAAGAAGAAATTATTAGG
 Srebrina     (1) TGTTTGAATGGTGGGGTTAAGGAAAGAAGAAGAAATTATTAGGTTTTTTTTGTGAGCAAGAAGAAATTATTAGG
   Avigea     (1) TGTTTGAATGGTGGGGTTAAGGAAAGAAGAAGAAATTATTAGGTTTTTTTTGTGAGCAAGAAGAAATTATTAGG
Romantika     (1) TGTTTGAATGGTGGGGTTAAGGAAAGAAGAAGAAATTATTAGGTTTTTTTTGTGAGCAAGAAGAAATTATTAGG
   Saikai     (1) TGTTTGAATGGTGGGGTTAAGGAAAGAAGAAGAAATTATTAGGTTTTTTTTGTGAGCAAGAAGAAATTATTAGG
Srebrina      (1) TGTTTGAATGGTGGGGTTAAGGAAAGAAGAAGAAATTATTAGGTTTTTTTTGTGAGCAAGAAGAAATTATTAGG
Consensus     (1) TGTTTGAATGGTGGGGTTAAGGAAAGAAGAAGAAATTATTAGGTTTTTTTTGTGAGCAAGAAGAAATTATTAGG

                74                                  Xba I                  146
       R5    (74) TGGGATATCTTGATTTAAGATTCCAACTAATCTCTAGATGATATTAATTAATTGTTGTTAATTATTTTTTCTT
   Galina    (74) TGGGATATCTTGATTTAAGATTCCAACCAATCTCTAGATGATATTAATTAATTGTTGTTAATTATTTTTTCTT
       S2    (74) TGGGATATCTTGATTTAAGATTCCAACTAATCTCTAGATGATATTAATTAATTGTTGTTAATTATTTTTTCTT
 Srebrina    (74) TGGGATATTTTGATTTAAGATTCCAACTAATCTCTAGATGATATTAATTAATTGTTGTTAATTATTTTTTCTT
   Avigea    (74) TGGGATATCTTGATTTAAGATTCCAACTAATCTTTAGATGATATTAATTAATTGTTGTTAATTATTTTTTCTT
Romantika    (74) TGGGATATCTTGATTTAAGATTCCAACTAATCTTTAGATGATATTAATTAATTGTTGTTAATTATTTTTTCTT
   Saikai    (74) TGGGATATCTTGATTTAAGATTCCAACTAATCTTTAGATGATATTAATTAATTGTTGTTAATTATTTTTTCTT
Srebrina     (74) TGGGATATCTTGATTTAAGATTCCAACTAATCTTTAGATGATATTAATTAATTGTTGTTAATTATTTTTTCTT
Consensus    (74) TGGGATATCTTGATTTAAGATTCCAACTAATCTTTAGATGATATTAATTAATTGTTGTTAATTATTTTTTCTT

                147
       R5   (147) AAATTTA
   Galina   (147) AAATTTA
       S2   (147) AAATTTA
 Srebrina   (147) AAATTTA
   Avigea   (147) AAATTTA
Romantika   (147) AAATTTA
   Saikai   (147) AAATTTA
 Srebrina   (147) AAATTTA
Consensus   (147) AAATTTA
```

Fig. 6. Sequences alignment between soybean cultivars and lines including *XbaI* site in the SW9-1 locus. Note: The *XbaI* digestion site is TCTAGA. The mutated nucleotide is marked in black and underlined. The sequences of cv. Srebrina demonstrated both positions C or T in the *XbaI* site detected in the heterozygous plant

S2/4/3/8 and R5/11/11/7 and in the cultivar Galina T mutation is not detected out of the 10 plants investigated (Figure 6). This result give as reason to conclude that the studied SNP is not related with the trait large seed in tested by us genotypes, but found in the Chinese genotypes by Li et al. (2019).

During the selection process the mutant T alleles detected in both parental genotypes Romantica and Saikai of line S2/4/3/8 probably are lost. It has to underline that the selection were followed strictly to the large-seeded phenotype. In the study of Li et al. (2019) was demonstrated that the frequency of the allele ss246792949T (called SW9-1T) is 70% for the large-seeded phenotype. They also report that selection efficiency of this allele is twice higher in the large-seeded than in the small seeded phenotype. Respectively the frequency of the allele ss264792929C (named SW9-1C) is 30% and is presented mainly in a small seeded. Recently it was published data presenting seed size traits association study using polymorphic SSR markers located on chromosome 8 (de los Reyes et al., 2022). The allele SW9-1 is located on chromosome 9. In the future work our intends are to apply SSR markers in order to make relation between seed size phenotype and support the soybean selection programs.

Conclusion

During two consecutive years F5 and F6 populations of seven elite soybean lines were investigated for the yield parameters. The lines were selected in purpose to develop a new soybean cultivar combining high yield and early maturity. Generally the increased yield of soybean is a result of growing land areas use, and a small part is related with the selection improvement of the soybean germplasm. That is the main difference between soybean and wheat and corn.

The collected data demonstrated that after crosses between Saikai as a mother genotype and Galina or Romantica as a father genotype we were able to obtain lines with high number of pods and seeds per plant. Same lines demonstrated also earlier ripening in comparison with the standard cultivar Avigea. The higher values of individual plant productivity and specific seed mass were achieved after crosses between Romantica and Srebrina. By application of CAPS marker for detection of SNP in the locus SW9-1 the mutant allele ss246792949T (called SW9-1T) were detected in the cultivars Romantica, Srebrina and Avigea with large-seeded phenotypes. Base on the results from the analysis of two elite soybean lines S2/4/3/8 and R5/11/11/7 we could conclude that probably this SNP is not related with large-seed phenotype in the studied genotypes in comparison with the Chinese genotypes. In our future work additional studies by application of SSR markers will be performed.

Acknowledgements

This study was supported by the Agricultural Academy, under the project "Physiological, molecular and phytopathological study of Bulgarian soybean varieties", number 29.

References

Aleksieva, A. (2015). Comparative evaluation of new soybean lines by economic properties. *In:* Scientific Session of Jubilee 90 years Experimental Station on Soybean, Pavlikeni, Bulgaria, 50-58.

Borowska, M. & Prusiński, J. (2021). Effect of soybean cultivars sowing dates on seed yield and its correlation with yield parameters. *Plant, Soil and Environment*, *67*(6), 360-366.

Cox, T. S., Kiang, Y. T., Gorman, M. B. & Rodgers, D. M. (1985). Relationship between coefficient of parentage and genetic similarity indices in soybean. *Crop Sci., 25,* 529–532.

de los Reyes, A. M., Santos, M. M. L., Ladia, Jr. V. A., Maghirang, R. G., Enicola, E. E. & Ocampo, E. T. M. (2022). Genotype profiling, population structure, and seed size traits association analyses using five polymorphic SSR markers in soybean [*Glycine max* (L.) Merr.] genotypes available in the Philippines. *Sci. Engg. J.* (the Official Journal of Philippine-American Academy of Science and Engineering), *166*(15), 2.

Hao, D. R., Cheng, H., Yin, Z. T., Cui, S. Y., Zhang, D., Wang, H. & Yu, D. Y. (2012). Identification of single nucleotide polymorphisms and haplotypes associated with yield and yield components in soybean (*Glycine max*) landraces across multiple environments. *Theor*. *Appl. Genet., 124*, 447–458.

Hudcovicova, M. & Kraic, J. (2003). Utilisation of SSRs for Characterisation of the Soybean (*Glycine max* (L.) Merr.) Genetic Resources. *Czech J. Genet. Plant Breed., 39*(4), 120–126.

Hipparagi, Y., Singh R., Choudhury, D. R. & Gupta, V. (2017). Genetic diversity and population structure analysis of Kala bhat (*Glycine max* (L.) Merrill) genotypes using SSR markers. *Hereditas, 154*, 9. DOI 10.1186/s41065-017-0030-8.

Keim, P., Beavis, W., Schupp, J. & Freestone, R. (1992). Evaluation of soybean RFLP marker diversity in adapted germ plasm. *Theor. Appl. Genet., 85*, 205–212.

Li, J., Zhao, J., Li, Y., Gao, Y., Hua, S., Nadeem, M., Sun, G., Zhang, W., Hou, J., Wang, X. & Qiu L. (2019). Identification of a novel seed size associated locus SW9-1 in soybean. *The Crop Journal, 7,* 548-559, doi.org/10.1016/j.cj.2018.12.010.

Liu, S., Zhang, M., Feng, F. & Tian, Z. (2020). Toward a "green revolution" for soybean. *Molecular Plant*, *13*(5), 688-697.

Mansur, L. M., Orf, J. H. & Lark, K. G. (1993). Determining the linkage of quantitative trait loci to RFLP markers using extreme phenotypes of recombinant inbreds of soybean (*Glycine max* (L). Merr.), *Theor. Appl. Genet., 86*, 914–918.

Messmer, M. M., Melchinger, A. E., Herrmann, R. G. & Boppermaier, J. (1993). Relationship among early European maize inbreds: II. Comparison of pedigree and RFLP data. *Crop Sci., 33*, 944–950.

Mian, M. A. R., Bailey, M. A., Tamulonis, J. P., Shipe, E. R., Carter, Jr., T. E., Parrott, W. A., Ashley, D. A., Hussey, R. S. & Boerma, H. R. (1996). Molecular markers associated with

seed weight in two soybean populations. *Theor. Appl. Genet., 93*, 1011–1016.

Naydenova, G. & Georgieva N. (2019). Study on seed yield components depending on the duration of vegetation period in soybean. *Bulg. J. Agric. Sci.*, *25*, 49-54.

Radkova, M. & Naydenova, G. (2022). Ecological and genotypic effects on traits harvest index and absolute seed weight in soybean. *Rastenievuni Nauki, 59*(2), 74-80 (Bg).

Shoemaker, R. C., Guffy, R. D., Lorenzen, L. L. & Specht, J. E. (1992). Molecular mapping of soybean: Map utilization. *Crop Sci., 32*, 1091–1098.

Song, Q., Jia, G., Zhu, Y., Grant, D., Nelson, T., Hwang, E. Y., Hyten, D. L. & Cregan, P. B. (2010). Abundance of SSR Motifs and Development of Candidate Polymorphic SSR Markers (BARCSOYSSR_1.0) in Soybean. *Genomics, Molecular Genetics & Biotechnology, 50*(5), 1950-1960. doi.org/10.2135/cropsci2009.10.0607.

Sun, Y. N., Pan, J. B., Shi, X. L., Du, X. Y., Wu, Q., Qi, Z. M., Jiang, H. W., Xin, D. W., Liu, C. Y., Hu, G. H. & Chen, Q. S. (2012). Multi-environment mapping and meta-analysis of 100-seed weight in soybean. *Mol. Biol. Rep., 39*, 9435–9443.

Wang, X. B., Li, Y. H., Zhang, H. W., Sun, G. L., Zhang, W. M. & Qiu, L. J. (2015). Evolution and association analysis of GmCYP78A10 gene with seed size/weight and pod number in soybean. *Mol. Biol. Rep., 42*, 489–496.

Received: **January, 04, 2024**; *Approved*: **March, 03, 2024;** *Published:* **December, 2024**