Journal of Sohag Agriscience (JSAS) 2020 (1):8-29 **ISSN 2357-0725**
Website: https://jsasblog.wordpress.com **Galal *et al* (2020)**

Genetic diversity of some soybean genotypes under drought stress using SRAP markers

Galal A. R. El-Sherbeny[1]; Abdelsabour G. A. Khaled[1]; Akram R. Morsy[2] and Mahmoud A. A. Abotaleb[2]

[1]Department of Genetics, Faculty of Agriculture, Sohag University, 82524 Sohag, Egypt.

[2]Department of Food Legumes, Shandaweel Island Station, Agricultural Research Center.

Keywords: molecular markers, variance, correlation, drought stress.

Abstract

Mean performance and genetic variability were assessed in four soybean genotypes and their F_1 crosses using phenotypic data and SRAP markers under drought stress conditions. The results showed that mean squares of G x E interaction were found to be highly significant for most studied traits. Also, the mean performances of four parents and their 6 crosses were variable from normal irrigation to drought stress condition. It could be observed that the parent genotype P_2 were relatively stress tolerant parent with the DSI value of 0.5. Using SRAP markers, a total of 49 bands were amplified, of which 24 bands (48.98%) were found polymorphic. Furthermore, the polymorphic band numbers ranged from 2 to 6 bands. The percentage of polymorphism (%P) ranged from 40% (ME2-EM10 primers) to 62% (ME7-EM6 primers) with an average of 48.85%. The SRAP marker ME5-EM1$_{590bp}$ was regarded probably as candidate marker which linked to plant height trait. Interestingly, three different markers (ME1-EM6$_{950bp}$, EM4-ME6$_{1000bp}$ and EM7-ME6$_{970bp}$) were regarded as candidate markers linked to number of branches per plant. The results showed highly significant and significant regressions (0.3249**, *P*= 0.013) and (45.04*, *P*= 0.053) on number of branches per plant and plant height traits, respectively. The UPGMA cluster analysis based on the SRAP markers and the means of morphological traits separated the soybean parental genotypes into two significantly different clusters. Finally, the correlation between the two markers is not significant (r = 0.565, *P*=0.932).

INTRODUCTION

Soybean (Glycine max (L.) Merrill) (2n = 2x = 40) is one of the most important oilseed crop in the world. It is grown on an estimated area of 125 million hectare, globally producing 341 million ton, it is supplies represent more than 60 % of the global demand of vegetable oil and protein, Approved by United States Department of Agriculture **(FAO stat 2013)**. Soybean, the 'golden bean', is consider to be a miracle crop in the world in terms of its use in human food and cattle feed, it is extraordinarily rich in protein about 40% and 21% oil on a dry weight basis **(Gopalan *et al.*, 1994).** Soybean protein contains many

essential amino acids, for both human and animals, mainly lysine, tryptophan, methionine and cysteine. Lecithin, extracted from soybean oil, is used for everything from pharmaceuticals to protective coatings. It is a natural emulsifier and lubricant. Soybean is the best and the cheapest source of protein for human beings and animals. Drought is an important abiotic factor limiting soybean production worldwide and drought alone accounts for about 40% crop loss. Irrigation and soil reclamation are not economically viable options for soybean production under drought conditions. The correlation and path analysis provide information on genetic association of yield contributing characters, which in turn are useful in developing breeding strategies. Association of characters influenced by a large number of genes is elaborated statistically by correlation coefficients. Genotypic correlation coefficient provides a measure of genotypic conjugation between characters.Morphological markers are greatly influenced by environmental factors. But, molecular markers are not influenced by environmental conditions for this reason these markers are decisive and more efficient for selection in breeding programs as well as for assessing genetic diversity amongst breeding materials. In relation to cost, easiness of use, consistency and repeatability of the results, sequence related amplified polymorphism (SRAP) is techniques based on polymerase chain reaction (PCR) used to assess genetic diversity. The SRAP was based on the amplification of open reading frame (ORF) developed from genome sequence data. Molecular markers provided excellent tools to estimate the genetic diversity **(Agarwal *et al.*, 2008).** SRAP markers used to study the genetic diversity among soybean genotypes **(Baloch *et al.*, 2010 and Sun *et al.*, 2013),** squash genotypes **(Ferriol *et al.*, 2003 and Inan *et al.*, 2012)**, sorghum **(Khaled *et al.*, (2019a)**, wheat **(Khaled *et al.*, (2019b)**, Brassica **(Li and Quiros, 2001)** and okra **(Robarts and Wolfe, 2014 and Yıldız *et al.*, 2016)**. The objectives of this investigation were directed to study the performance of four soybean genotypes and their F_1 crosses under drought stress, estimate of genetic variability using SRAP molecular markers and identify the association between morphological and molecular markers.

MATERIALS AND METHODS

Genetic Materials:

Field experiments were conducted at the Experimental Farm of Agricultural Research Station in Shandaweel Island, Sohag Governorate, Egypt in two consecutive seasons 2014-2015 and 2015-2016. Four soybean genotypes (Table 1) were used in this study. The extraction of genomic DNA and PCR procedures were conducted at molecular genetics lab., Department of Genetics, Faculty of Agriculture, Sohag University.

Table 1: List and data of four soybean cultivars studied under normal and drought stress conditions.

Genotypes		Pedigree	Country of origin	characters
PI 416937	(P_1)	Japan	Japan	Resistant
Giza 111	(P_2)	Crawford x Celest	Field Crops Research Institute, Giza, Egypt	Sensitive
H10 L228	(P_3)	N92-83 x Giza111		
H6 L1	(P_4)	Giza 83 x Ware		

Field experiment

Pollen of male parents was collected from newly opened flowers. Pollen was place on the stigma of the female parent by tapping it gently with the anthers. When pollination was completed, the flower was identified with a tag and the other flowering buds were removed from the same cluster. Hybridization was more successful in the early morning**.**

Half-diallel cross mating design among four soybean genotypes was made to produce six crosses in 2015 growing season. In the second season of 2016, the seeds of four genotypes and their six F_1 crosses were sown in field, in two experiments. The first experiment was grown under supplemental water applied regularly as recommended (Normal, N) while, the second experiment received half of the number of irrigation compared to first experiment (drought stress, D). Each experiment was evaluated in a randomized complete block design (RCBD) with three replications. Each block contains 10 plots. Each plot consisted of one row. Each row was 3m long and 30 cm between rows. Plants were spaced by 15 cm. within row. Planting and transplanting dates were 20th may in 2015 and 25th may 2016, respectively. The experimental field soil was sandy clay loam in texture. Normal agronomic practices of growing soybean were carried out until harvest. All recommended cultural practice was applied under normal conditions (every 10 days) and drought stress (every 20 days). Five plants were selected randomly from each genotype

at each treatment for recording the observations. Average value for each character was computed from these plants separately for each genotype and observations were recorded on; days to 50% flowering, days to 90 % maturity, plant height (cm), number of branches/plant, number of pods/plant, seed yield/plant (gm) and 100 seed weight (gm).

Analysis of variance:

Data were subjected to general analysis of variance for Randomized Complete Block Design (RCBD) according to **Steel and Torrie (1980)**. Mean squares of genotypes and replications for all studied traits were tested for significance according to the F-test. The form of analysis of variation (S.O.V) was outlined by **Cochran and Cox (1957)**. The appropriate variances and covariances were used to calculate phenotypic and genotypic correlation coefficients **(Johnson, *et al.* 1955)**. Significance of the various correlation coefficients was tested from the statistical table of correlation coefficients at 1 and 5 per cent level of significance **(Snedecor and Cochran, 1967).**

Drought susceptibility index (DSI):

Drought Susceptibility Index was computed according to **Fischer and Maurer (1978)** equation. Genotypes with average susceptibility or resistance to drought have an "S" value of 1.0. Less than 1.0 indicate less susceptibility and greater resistance to drought. While, a value of S=0 indicates maximum possible drought resistance (no effect of drought on yield) **Fischer and Maurer (1978)**.

DNA extraction and PCR procedure:

Fresh young leaves of the four parental soybean genotypes were harvested and immediately grinded in extraction buffer using cetyltrimethyl ammonium bromide (CTAB) protocol as described by **Porebski *et al.* (1997)**, at molecular genetic Laboratory of genetics Department, Faculty of Agriculture, Sohag University. For each genotype, 0.2 gm of grinded leaf tissue was suspended in 2 ml of extraction buffer (20 mM of EDTA, 0.1 M of Tris-HCL, 1.4 of Nacl, 2% CTAB, 1% of PVP). The DNA pellet was then suspended in 100 µl of TE buffer. Genomic DNA was diluted 10-folds in water prior to 35 cycles of PCR amplification. The PCR assays were performed in a 20µl volume containing 0.2µl of Go Taq polymerase, 3.5µl of primer (8 pmol), 4µl 5X green buffer, 2µl $MgCl_2$, 2µl dNTPs (2.5mM), 5.3µl of free nuclease water and 3µl (150-200 ng) of genomic

DNA templates. The thermal Cycler 96-Labmet (USA) was programmed as following: 5 cycles comprising 1 min at 94ºC (denaturation), 1 min at 35ºC (annealing) and 1 min 30 sec at 72ºC (elongation), in the following 35 cycles, denaturation at 94ºC for 1 min, annealing at 49ºC to 56ºC for 1 min and elongation at 72ºC for 2 min, ending with an extending step for 10 min at 72ºC. The amplified products were electrophoresed in a 1% agarose gel stained with 0.2 µl ethidium bromide. The amplified fragments were visualized and photographed using UVP Bio Doc-It imaging system (USA) **Sambrook *et al.* (1989)**. SRAP technique was conducted using 15 (forward and reverse) primer combinations (Table 2).

Data analysis of molecular markers:

The DNA banding patterns generated by SRAP were analyzed by computer program Gene Profiler software (version 4.03). The presence (1) or absence (0) of each band was recorded for each genotype for all studied primers. Genetic distance was estimated according to **Jaccard (1908)**.To measure the informativeness of the SRAP technique in differentiating among genotypes, the polymorphic information content (PIC) was calculated according to the formula of **Ghislain *et al.* (1999)**as PIC= 1- $[(p)^2 + (q)^2]$ where p is the frequency of allele band present and q is frequency of allele band absent across the studied genotypes. The marker index (MI) was calculated for each primer as MI = PIC x $\eta\beta$, where PIC is the mean PIC value, η the number of bands, and β is the proportion of polymorphism **Powell *et al.* (1996)**. Analysis of variance (ANOVA) was conducted using the 0–1 data. The association analysis was conducted using simple linear regression. For this, data on individual phenotypic traits were regressed on whole 0-1 binary marker data for each individual phenotypic marker using Excel programme. The coefficient of determination (R^2) was calculated as R^2 = 1- (SSE/SST), where SSE is the sum of squares of error and SST is the total sum of squares.

Dendrograms construction:

The genetic similarities among the studied genotypes were computed and UPGMA-dendogram was performed according to Jaccard's coefficient **(Jaccard, 1908)** using the computational package MVSP version 3.1. A cophenetic matrix was derived from each matrix to test goodness of fit of the clusters by comparing the matrices using Mantel test **(Mantel,**

1967). Finally, the correlation between SRAP and means of morphological traits was calculated using NTSYS-pc version 2.20 **(Rohlf, 2000).**

Table 2. SRAP primers, sequences and Tm ºC used in this study.

Primers name	Primer Sequence (5' – 3')	TmºC
ME-1(F)	TGAGTCCAAACCGGATA	49
ME-2(F)	TGAGTCCAAACCGGAGC	54
ME-4(F)	TGAGTCCAAACCGGACC	56
ME-5(F)	TGAGTCCAAACCGGAAG	52
ME-7(F)	TGAGTCCAAACCGGACG	54
EM-1(R)	GACTGGGTACGAATTAAT	50
EM-3(R)	GACTGCGTACGAATTGAC	50
EM-6(R)	GACTGCGTACGAATTGCA	53
EM-9(R)	GACTGGGTACGAATTCAC	52
EM-10(R)	GACTGGGTACGAATTCCA	53

Results and Discussion

Analysis of genotypic variation:

Earliness and vegetative traits:

The analysis of variance showed the differences between mean squares of environments were highly significant for days to flowering, days to 50% maturity, plant height and number of branches per plant (Table 3). Also, mean squares of genotypes under each environment and the combined data were found highly significant. Moreover, mean squares due to G×E were also highly significant for these traits.

Yield component traits:

The analysis of variance for 100 seed weight and Seed yield per plant revealed highly significant difference between both environments. While it indicated no significant for number of pods per plant trait (Table 3). Mean squares of genotypes were found to be highly significant for genotypes under each of environment and the combined data. Generally, the results of this study showed that mean squares of G x E interaction were found to be highly significant for all studied traits. This finding suggested a differential response of the genotypes from environment to another. Similar results were obtained by **Yadav (2005); Karnwal and Singh (2009); Mahbub *et al.* (2015)** and **Ghiday *et al.* (2016).**

Mean Performance:

Earliness and vegetative traits:

The results cleared that the mean performance of all studied traits for the four parents and their sex F_1 hybrids were varied from normal to drought stress conditions (Table 4). The genotype (P_3) was found to be the earliest parent with the mean value of 32.7 days under drought conditions, while genotypes (P_1) and (P_2) were the latest parents with a mean of 40.5 and 39.1 days under drought stress conditions. Concerning F_1 hybrids, the mean performance of earliness were narrower than their parents ones under both environments and their combined data. The results showed that the best combination for earliness was (P_2xP_3) and (P_3xP_4). The earliest parent in maturity was genotype P_3 under both environments and their combined data. However, the cross (P_3 xP_4) recorded the lowest in number of days to maturity at the two environments and their combined data. The data pointed out that the parent (P_3) is considered the tallest variety under both normal condition (98.30 cm), stress condition (87.90 cm) and (94.00 cm) for combined analysis data. On the other hand, the parental (P_1) is considered as the shortest parental under both normal and drought conditions for combined analysis data (61.00 cm, 47.70 cm and 54.30 cm), respectively. Regarding to the F_1 hybrids, the cross combination (P_3xP_4) is considered the tallest cross, while the hybrid (P_1xP_3) is considered the shortest hybrid. The results revealed that the genotype P_2 (3.77) had the highest value for no. of branches/plant under normal condition. Also, the cross (P_2 x P_3) had the highest mean values of no. of branches/plant under normal (4.67), stress irrigation (4.48) and in combined (4.57).

Yield component traits:

The genotype (P_3) gave the highest mean value of number of pods per plant (Table 4) under both conditions and the combined data (129.60,111.80 and 120.70 pods/plant) respectively, while the F_1 hybrid (P_2xP_3) had the highest mean values under normal condition (116.40 pods/plant) followed by (P_2xP_3) had the highest mean values under stress condition (118.90 pods/plant) and (P_3xP_4) had the highest mean values under the combined of both environments (111.80 pods/plant). The results of 100 seed weight trait (Table 5) showed that (P_4) had the highest mean value of 22.30gm under normal condition and (P_1) had the highest mean values of 22.80gm and 22.50gm under stress environment and the combined data, respectively. The cross

(P_1xP_2) had the highest mean values of 100 seed weight (24.80gm), (23.80gm) and (24.30gm) under normal and stress treatment and combined data respectively. The result of seed yield per plant (Table 5) illustrated that (P_3) had the highest mean values (198.3gm, 158.4gm and 178.3gm, respectively) under normal, stress treatment and the combined both environments respectively. On the other hand, the crosses of (P_2xP_3) yielded more seeds per plant (172.6gm) in comparison with the other crosses under normal irrigation. Under stress treatment, (P_1xP_3) gave the highest mean value (170.0gm). In addition to (P_3xP_4) recorded the highest mean value The results revealed that the genotype P_2 (3.77) had the highest value for no. of branches/plant under normal condition. Also, the cross (P_2 x P_3) had the highest mean values of no. of branches/plant under normal (4.67), stress irrigation (4.48) and in combined (4.57).

Yield component traits:

The genotype (P_3) gave the highest mean value of number of pods per plant (Table 4) under both conditions and the combined data (129.60,111.80 and 120.70 pods/plant) respectively, while the F_1 hybrid (P_2xP_3) had the highest mean values under normal condition (116.40 pods/plant) followed by (P_2xP_3) had the highest mean values under stress condition (118.90 pods/plant) and (P_3xP_4) had the highest mean values under the combined of both environments (111.80 pods/plant).

The results of 100 seed weight trait (Table 5) showed that (P_4) had the highest mean value of 22.30gm under normal condition and (P_1) had the highest mean values of 22.80gm and 22.50gm under stress environment and the combined data, respectively. The cross (P_1xP_2) had the highest mean values of 100 seed weight (24.80gm), (166.8gm) in the combined both environments.

Reduction due to drought stress (R %): The results of reduction due to drought stress for all studied traits are given in Table 5. The results showed that the mean performances of four parents and their 6 crosses were variable from normal irrigation to drought stress condition. It could be concluded that the drought stress caused reduction about 9.37, 4.44, 12.9, 6.71, 4.34, 4.79% and 9.55% in the parental varieties and hybrids average for days to 50% flowering, days to 50% maturity, plant height, no. of branches per plant, no. of pods per plant, 100 seed weight and seed yield per plant, respectively.

(23.80gm) and (24.30gm) under normal and stress treatment and combined data respectively.

The result of seed yield per plant (Table 5) illustrated that (P_3) had the highest mean values (198.3gm, 158.4gm and 178.3gm, respectively) under normal, stress treatment and the combined both environments respectively. On the other hand, the crosses of (P_2xP_3) yielded more seeds per plant (172.6gm) in comparison with the other crosses under normal irrigation. Under stress treatment, (P_1xP_3) gave the highest mean value (170.0gm). In addition to (P_3xP_4) recorded the highest mean value (166.8gm) in the combined both environments.

Reduction due to drought stress (R %):

The results of reduction due to drought stress for all studied traits are given in Table 5. The results showed that the mean performances of four parents and their 6 crosses were variable from normal irrigation to drought stress condition. It could be concluded that the drought stress caused reduction about 9.37, 4.44, 12.9, 6.71, 4.34, 4.79% and 9.55% in the parental varieties and hybrids average for days to 50% flowering, days to 50% maturity, plant height, no. of branches per plant, no. of pods per plant, 100 seed weight and seed yield per plant, respectively.

Drought Susceptibility Index (DSI):

The estimates of drought stress susceptibility index (DSI) based on seed yield/plant for four parents and their six F_1 crosses showed in Table (5). It could be observed that the parent genotype P_2 were relatively stress tolerant parent with the DSI value of 0.5. Concerning the F_1 hybrids, the cross combinations (P_1xP_2), (P_1xP_3) and (P_3xP_4) were relatively tolerant to drought stress with the DSI values of 0.8, 0.7 and 0.4, respectively. These results indicated that the tolerant parent was P_2 transmitted their genes controlling drought tolerance to their hybrids. Consequently, these crosses could be considered as promising populations for isolating useful segregates to be cultivated under drought stress. Similar results were found by **Frederick *et al*. (2001)**; **Wang *et al*. (2009); Fenta *et al.* (2012); Devi and Sinclair (2013) and Pathan *et al.* (2014).**

Table 3: Analysis of variances and mean squares of the four parents and their F_1 hybrids for the studied traits under normal (N), drought (D) conditions and combined over environments (C).

S.V	D.F		Mean squares											
			FD			MD			P.H			NB/P		
	S	C	N	D	C	N	D	C	N	D	C	N	D	C
E	---	1	---	---	214.5**	---	---	496.5**	---	---	1672.6**	---	---	0.97**
R	2	---	0.09	0.17	---	2.44	0.82	---	1.27	0.06	---	0.10	0.01	---
R /E	---	4	---	---	0.13	---	---	1.63	---	---	0.66	---	---	0.06
G	9	9	50.1**	25.8**	72.8**	195.1**	177.1**	365.2**	703.5**	622.2**	1307.3**	1.02**	1.23**	1.84**
G x E	---	9	---	---	3.12**	---	---	7.01**	---	---	18.5**	---	---	0.41**
Error	18	36	0.33	0.15	0.24	1.01	0.29	0.65	0.87	2.04	1.45	0.12	0.04	0.08

S.V	D.F		Mean squares								
			NP/P			100 SW			S Y/P		
	S	C	N	D	C	N	D	C	N	D	C
E	---	1	---	---	266.8	---	---	12.9**	---	---	3129.7**
R	2	---	118.5*	100.2*	---	0.23	0.07	---	315.7**	28.0	---
R /E	---	4	---	---	109.4	---	---	0.15	---	---	171.9
G	9	9	1001.7**	1206.6**	2024.9**	23.5**	26.6**	49.0**	2129.5**	2196.5**	3987.4**
G x E	---	9	---	---	183.5**	---	---	1.09**	---	---	338.6**
Error	18	36	27.2	28.2	27.7	0.10	0.14	0.12	48.3	46.2	47.3

*, ** Significant at 5% and 1% levels of probability, respectively

E: Environment; R: Replication; G: Genotypes

FD, Days to flowering; MD, Days to maturity; PH, Plant height, NP/P, number of pods per plant; 100 SW, 100 seed weight and S Y/P, Seed yield per plant.

Table 4: Mean performance of parents and F_1's for some studied traits under normal (N), drought (D) conditions and combined over environments (C).

G.	F.D			M.D			P.H			No. of B/P		
	N	D	C	N	D	C	N	D	C	N	D	C
P1	47.0	40.5	43.8	144.3**	139.9**	142.1	61.0	47.7	54.3	3.11	3.17	3.14
P2	44.4	39.1	41.7	127.4	120.8	124.1	95.4**	81.9**	88.7	3.77	3.23	3.50
P3	35.4**	32.7**	34.1	124.1	115.1	119.6	98.3**	87.9**	94.0	3.17	2.70	2.93
P4	37.5**	33.7**	35.6	133.4**	126.5**	129.9	85.5**	75.9**	80.7	3.65	3.60	3.63
P1*P2	43.0	37.9	40.5	138.7**	131.3**	135.0	72.0	57.7	64.9	4.43*	4.37**	4.40
P1*P3	42.9	39.3	41.1	129.5	122.4	126.0	62.4	60.2	61.3	4.60*	3.63	4.12
P1*P4	40.4**	37.0**	38.7	131.9**	125.1**	128.5	65.9	57.2	61.5	3.70	2.63	3.17
P2*P3	35.2**	33.1**	34.2	126.6	123.2	124.9	91.4**	80.5**	85.9	4.67*	4.48**	4.57
P2*P4	40.9**	38.3	39.6	122.9	120.0	121.5	86.4**	76.2**	81.3	3.32	3.37	3.34
P3*P4	36.2**	33.8**	35.0	116.3	113.4	114.9	100.9**	88.2**	94.6	3.43	4.13**	3.78
mean	40.3	36.5	_	129.5	123.8	_	81.9	71.3	_	3.79	3.53	_
L.S.D 5%	0.99	0.68	0.81	1.73	0.92	1.33	1.60	2.45	1.99	0.60	0.36	0.47
L.S.D 1%	1.35		0.93	1.09		2.36	1.27		1.78	2.20		3.36

Table 5: Mean performance of parents and F_1's for some studied traits under normal (N), drought (D) conditions and combined over environments (C).

G	No. of P/P			100 S.W			S.Y/P			DSI
	N	D	C	N	D	C	N	D	C	
P_1	103.0	97.0	100.0	22.2**	22.8**	22.5	97.3	87.8	92.5	1.0
P_2	89.3	87.7	88.5	18.2	16.2	17.2	149.6	142.8	146.2	0.5
P_3	129.6**	111.8**	120.7	19.5	17.8	18.6	198.3**	158.4**	178.3	2.1
P_4	72.7	63.2	68.0	22.3**	20.7**	21.5	141.0	112.6	126.8	2.7
P_1*P_2	83.3	86.3	84.8	24.8**	23.8**	24.3	141.9	131.7	136.8	0.8
P_1*P_3	102.1	118.9**	110.5	17.8	16.0	16.9	170.8**	160.0	165.4	0.7
P_1*P_4	93.8	79.4	86.6	16.6	15.8	16.2	153.4	126.6	140.0	1.8
P_2*P_3	116.4**	103.0*	109.7	18.5	18.1	18.3	172.6**	154.5**	163.6	1.1
P_2*P_4	73.5	65.0	69.3	16.7	16.7	16.7	132.7	114.2	123.5	1.4
P_3*P_4	107.2*	116.3**	111.8	17.0	16.1	16.6	169.6**	164.0*	166.8	0.4
mean	**97.1**	**92.9**	**_**	**19.3**	**18.4**	**_**	**151.2**	**136.8**	**_**	
L.S.D 5%	8.95	9.11	8.69	0.55	0.65	0.58	11.9	11.7	11.3	
L.S.D 1%	12.3	12.5	11.6	0.76	0.89	0.78	16.3	16.0	15.2	
Reduction	**4.34**		**-**	**4.79**		**-**	**9.55**		**-**	

SRAP Markers:

Detecting of DNA Polymorphism:

Sequence Related Amplified Polymorphism (SRAP) molecular markers are based on two primers amplification, which preferentially amplifies open reading frames (ORFs) or coding regions using polymerase chain reaction (PCR). SRAPs are molecular markers which could provide high polymorphism and plentiful information to assess the genetic diversity. For this purpose, in this study, fifteen pairs (forward and reverse) of SRAP primer combinations were screened among the four soybean parental genotypes, and 6 pairs of them were polymorphic (Table 6 and Figure 1). A total of 49 bands were amplified, of which 24 bands (48.98%) were found polymorphic (Table 6). The total number of bands varied from 5 (ME2-EM10) to 13 (ME1-EM6). Furthermore, the polymorphic band numbers ranged from 2 to 6 bands. The mean number of total bands and polymorphic bands were 8.17 and 4 per primer, respectively. The percentage of polymorphism (%P) ranged from 40% (ME2-EM10) to 62% (ME7-EM6) with an average of 48.85%. The molecular weight of bands ranged from 335 bp to 1250 bp, generated by ME5-EM1 and ME1-EM6 primer combinations, respectively.

The results of present study are in agreement with those obtained by **Sun *et al.*, (2013)** showed that the 18 SRAP primer combinations detected a total of 90 polymorphism bands (5 per primer combination) with a mean of gene diversity of 0.918. In this regard, **Baloch *et al.*, (2010)** used thirty-four SRAP primer combinations which produced a total of 155 scorable bands, with an average of 4.66 bands per primer combination, of which 26 (17%) were polymorphic. The total number of amplified bands was between 2 (Me1Em9, Me2Em2, Me3Em2, Me3Em4, and Me7Em1) and 8 (Me1Em2 and Me2Em1); the number of polymorphic bands ranged from 0 to 3.

Using ISSR markers **Jain *et al.*, (2017)** obtained 177 polymorphic bands (97.25% polymorphism) for twenty four soybean genotypes. Additionally, the number of amplified bands varied from 3 (UBC-872 and UBC-878) to 17 (UBC-814) with an average of 9.57 per primer.

Regarding other economical crops, **Khaled *et al.*, (2019a)** documented that out of 82 bands, 39 bands (47.56%) were polymorphic applying SRAP markers for seven sorghum

genotypes. They showed that the average of %P was 47.09%. The mean number of total bands and polymorphic bands were 6.83 and 3.54 per primer, respectively. Using SRAP markers to detect the genetic variability between 8 wheat genotypes **Khaled *et al.*, (2019b)** observed a total of 95 bands, of which 64 bands (62.59%) were polymorphic. The total number of bands ranged from 2 to 12. The average of P% was 62.59%. The mean number of total bands and polymorphic bands were 7.92 and 5.3 per primer, respectively. The size of polymorphic bands ranged from 200 bp to 1750 bp.

Table 6: Primers used for SRAP marker, total number of fragment detected by each pair of primers, %P, PIC, MI and fragments sizes.

Primer Combinations	Amplified bands		%P	PIC	MI	Fragments size	
	Bands number	Polymorphic bands				Larger (bp)	Smaller (bp)
ME1-EM6	13	6	46.15	0.19	1.14	1250	430
ME5-EM1	9	4	44.44	0.18	0.72	980	335
ME4-EM6	8	4	50.00	0.21	0.84	1200	505
ME7-EM6	8	5	62.50	0.25	1.25	990	375
ME1-EM3	6	3	50.00	0.21	0.63	655	400
ME2-EM10	5	2	40.00	0.15	0.30	700	430
Total	**49**	**24**					
Mean	**8.17**	**4.00**	**48.85**	**0.20**	**0.81**		

%P: Percentage of polymorphism,
PIC: Polymorphic information content,
MI: Marker index.

Single marker analysis:

The present study involved a set of four soybean genotypes, exhibiting moderate to high genetic variability for the phenotypic traits included in this work. Using simple linear regression method, a total of 24 polymorphic molecular markers were identified; four of them were significantly associated with two different traits.

Results in (Table 7 and Figure 1) showed that the SRAP marker **ME5-EM1$_{590bp}$** was regarded probably as candidate marker which linked to Plant height (PH) trait. Interestingly, three different markers (**ME1-EM6$_{950bp}$, EM4-ME6$_{1000bp}$** and **EM7-ME6$_{970bp}$**) were regarded as candidate markers linked to one trait (Number of branches per plant, NB/P).The results showed highly significant and significant regressions (0.3249**, *P*= 0.013) and (45.04*, *P*= 0.053) on number of branches per plant and plant height traits, respectively. With SRAPs, some fragments were uniquely amplified in single genotypes. These markers are of great interest in optimal management of germplasm collections, as they facilitate the identification of genotypes and/or trait of interest. In bread wheat experiment, **Khaled and Hamam (2015)** reported that the SRAP marker ME7-EM6$_{420bp}$ was regarded as candidate marker linked to number of kernel/spike. Also, **Meng *et al.*, (2012)** reported that the primer pairs ME6/OD3, ME8/EM14, ME9/OD3, ME21/EM8 and ME21/EM18 showed correlation in the bulk for round fruit cucumber shape.

Regarding other molecular markers, **Ntuli *et al.*, (2015)** found unique alleles specific for CPSP population which detected by SSR marker: CMTP9 (151 bp); CMTP132 (134 bp) and PKCT111 (200 bp and 202 bp). Also, they showed that RAPD primers CB9, CB21 and CB15 identified unique bands of 400 bp, 900 bp and 200 bp, respectively in studied populations**. El-Sherbeny *et al.*, (2018)** identified an ISSR marker (UBC-825$_{490bp}$) which may be linked to number of pod per plant trait in Okra.

Generally, the molecular markers that respond most consistently and to the greatest extent in the target environment are the prime candidates for marker-assisted selection (MAS). Therefore SRAPs identified during the present study need to be subjected to validation and/or functional analysis of respective traits. In addition, at least one of the markers identified in the present investigation would be validated and used for MAS.

Table 7: Details of variance (ANOVA) involving coefficient of determination (R^2) for traits using 24 SRAP polymorphic bands.

Marker	Traits	S.V	df	SS	MS	R^2	*P-value*
ME1-EM6$_{950bp}$	NO. B/P	Genotypes	1	0.3249	0.3249**	0.974	**0.013**
EM4-ME6$_{1000bp}$		Error	2	0.009	0.0045		
EM7-ME6$_{970bp}$		Total	3	0.3339	771.20		
ME5-EM1$_{590bp}$	PH	Genotypes	1	771.20	45.04*	0.989	**0.053**
		Error	2	90.09			
		Total	3	861.29			

S.V: Source of variance**,**
d.f: Degrees of freedom**,**
S.S: Sum of squares**,**
M.S: Mean squares**,**
R^2: Coefficient of determination**.**

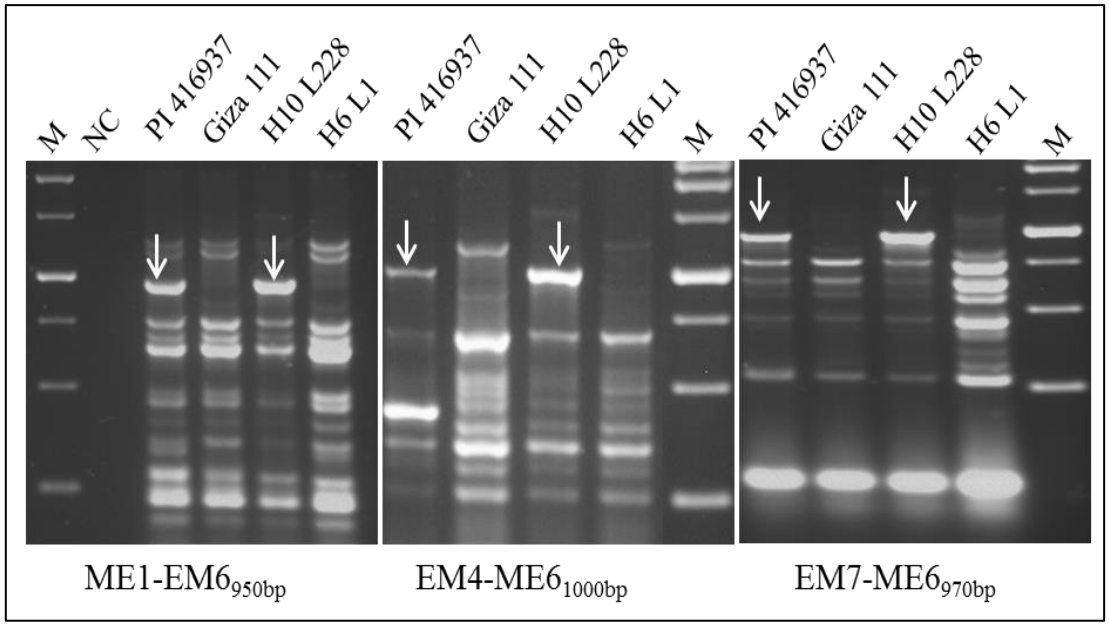

Figure 1: Gel profiles of SRAPs shows the polymorphism and the markers identified in this study.

PIC and MI analysis:

The polymorphism information content (PIC) index has been used extensively in many genetic diversity studies **(Tatikonda *et al.* 2009; Thudi *et al*. 2010)**. Moreover, the PIC value of markers indicates the usefulness of DNA markers for gene mapping, molecular breeding and germplasm evaluation **(Peng and Lapitan, 2005)**. In this study the PIC values for SRAP primer combinations varied from 0.15 (ME2-EM10) to 0.25 (ME7-EM6) with an average of 0.20 (Table 6). The mean PIC value of present study is lower than this (0.344) obtained by **Jain *et al*., (2017)**.

Marker index (MI)

May provide a convenient estimate of marker utility. In this work, the marker index values were from 0.30 (ME2-EM10) to 1.25 (ME7-EM6) with an average of 0.81 per primer combinations.

Cluster analysis:

A cluster analysis realized using Jaccard's coefficient **Jaccard, (1908)** for the data of SRAP molecular marker, revealed percent of similarity ranging from 65.20 (P_1 and P_2) to 82.10 (P_2 and P_4) (Table 8,below diagonal).The UPGMA cluster analysis based on the SRAP markers separated the soybean parental genotypes into two significantly different clusters. The first cluster was with **H6 L1** (P_4) and **Giza 111** (P_2) which branched at 82.10% percent of similarity (Figure 2 (A)).The parental genotypes **H10 L228** (P_3) branched at 71.70% with **I416937** (P_1) in the second cluster. **Powell *et al.* (1996)** demonstrated that several factors might affect the estimates of genetic relationships between individuals i.e., number of markers used, distribution of markers in the genome and the nature of evolutionary mechanisms underlying the variation measured.

Based on the means of morphological and agronomical studied traits, the cluster analysis revealed similarity percent ranging from 81.64% (P_1 and P_3) to 94.90% (P_2 and P_4) (Table 8, above diagonal). The dendogram divided the soybean parental genotypes into two different clusters (Figure 2 (B)). The first cluster subdivided into two subclusters. The first subcluster was with genotypes **H6 L1(P_4)** and **Giza 111(P_2)** which branched at 94.90% percent of similarity. The genotype **H10 L228 (P_3)** was in the second subcluster which branched at 88.92% percent of similarity with the first subcluster. The second cluster contains **PI 416937 (P_1)** branched at 85.66% percent of similarity with the first cluster. **Baloch *et al.*, (2010)** showed an average of Jaccard genetic similarity among soybean cultivars and advanced lines varied from 0.911 (61/21-Yeşilsoy) to 1.000 (S4240-16/21), with an average of 0.959.

Table 8: Similarity matrix for four soybean parental genotypes obtained from RAPD analysis (below diagonal), and similarity matrix obtained using phenotypic data (above diagonal).

Genotypes	(P_1) PI 416937	(P_2) Giza 111	(P_3) H10 L228	(P_4) H6 L1
(P_1) PI 416937	-	87.62	81.64	87.73
(P_2) Giza 111	65.20	-	90.66	94.90
(P_3) H10 L228	71.70	72.50	-	87.17
(P_4) H6 L1	65.30	82.10	68.20	-

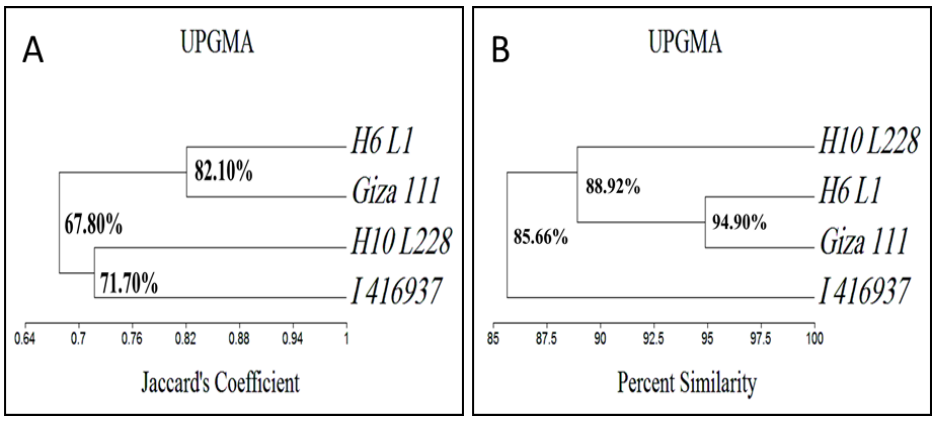

Figure 2: UPGMA-Dendrograms of genetic similarities of four soybean genotypes using SRAP data based on Jaccard's coefficient (A) and based on agronomic traits (B.

Combined molecular markers and morphological markers:

The correlation (*r*) and the Mantel test statistic (Z) were calculated to measure the degree of relationship between the similarity matrices obtained with SRAP molecular markers and phenotypic data (Figure 2). Results showed that the correlation was positive but insignificant (r = 0.565, *P*=0.932). Contrary, **Athanasios *et al.* (2009)** showed that a highly significant (r = 0.52; *P*= 0.002) correlation was found among phenotypic and RAPD markers. **Khaled *et al.* (2019b)** obtained a positive but non-significant correlation (r = 0.03, *P> 0.05*) between SRAP markers and means of wheat morphological traits. In the same direction, **Pandey *et al.* (2008)** showed that DNA markers are preferable to morphological ones because they relate variability directly at genetic level and provide reliable and enormous data that permit a reproducible estimate of genetic diversity in the germplasm. Molecular evaluation was more favorable than phenotypic evaluation because it had more markers and represented neutral traits of simple inheritance **(Sensoy *et al.*, 2007).**

Finally, the optimal strategies of the breeding system require extensive knowledge of the breeding materials employed. Results presented here will be useful to understand the current status of genetic diversity between soybean genotypes. Genetic markers like SRAPs may accurately assay the degree of genetic change between two genomes, but they may not necessarily reflect the divergence in terms of changes in traits of agronomic importance.

REFERENCES

Agarwal, M.; N. Shrivastava and H. Padh (2008): Advances in molecular marker techniques and their applications in plant sciences.Plant Cell Reports, 27(4):617-631.

Athanasios, L.T.; O. Koutita; A. Anastasiadou; G.N. Skaracis; E. Traka-Mavrona and M. Koutsika-Sotiriou (2009): Description and analysis of genetic diversity among squash accessions. Brazilian Archives of Biology and Technology, 52(2):271-283.

Baloch, F. S.; C. Kurt; H. ArIoglu and H. Ozkan (2010): Assaying of diversity among soybean (Glycin max (L.) Merr.) and peanut (Arachishypogaea L.) genotypes at DNA level. Turk J. Agric. For, 34 : 285-301.

Cochran, W.G. and G.M. Cox (1957): Experimental Designs. Wiley, New York.

Devi, M. J. and T. R. Sinclair (2013): Fixation drought tolerance of the slow-wilting soybean PI 471938. Crop Sci., 53: 2072–2078.

El-Sherbeny, G.A.R.; A.G.A. Khaled; H.A. Obiadalla-Ali and A.Y.M. Ahmed (2018): ISSR markers linked to agronomic traits in okra. International Journal of Modern Agriculture, 7(1):1-7.

FAOstat:http://faostat.fao.org/site/567/DesktopDefault.aspx?PageID=567#ancor. Updated 16.1.2013.

Fenta, B.A.; S.P. Driscoll;K. J. Kunert and C.H. Foyer (2012): Characterization of drought-tolerance traits in nodulated soya beans: the importance of maintaining photosynthesis and shoot biomass under drought-induced limitations on nitrogen metabolism. J. Agron. Crop Sci., 198, 92–103.

Ferriol, M.; B. Pico, and F. Nuez, (2003): Genetic diversity of some accessions of Cucurbita maxima from Spain using RAPD and SRAP markers. Genet. Resour. Crop Evol, 50:227-238.

Fischer, R.A. and R. Maurer (1978): Drought resistance in spring wheat cultivars. I-Grain yield responses. Australian J. of Agricultural Research, 29:897-912.

Frederick, J.R.; C.R. Camp and P.J. Bauer (2001): Drought stress effects on branch and main stem seed yield and yield components of determinate soybean. Crop Sci., 41: 759–763.

Ghiday, T.; A. Amogne ; G. Tefera and M.Malede (2016): Heritability, Genetic Advance and Path Coefficient Analysis for Grain Yield and its Component Characters in Soybean [Glycine max (L.) Merril], International Journal of Research Studies in Agricultural Sciences (IJRSAS), 3(5):1-11.

Ghislain, M.; D. Zhang; D. Fajardo; Z. Hanuman and R. Hijmans (1999): Marker assisted sampling of the cultivated Andean potato (Solanum phureja) collection using RAPD markers. Genetic Resources and Crop Evolution, 46: 547-555.

Gopalan, C.; B.V. Ramashastry and S.C. Balasubramanian (1994): Nutritive Value of Indian Foods, Indian Council of Medical Research, pp. 24-26.

Inan, N. ; M. Yildiz; S. Sensoy; S. Kafkas and K. Abak (2012): Efficacy of ISSR and SRAP techniques for molecular characterization of some cucurbita genotypes including naked (hull-less) seed pumpkin.The Journal of Animal & Plant Sciences, 22(1): 126-136.

Jaccard, P. (1908): Nouvelles recherché sur la distribution florale. Bulletin de la Societe Vandoise des Sciences Naturelles,44: 223-270.

Jain, R. K.; A. JoshiI; D. Jain; D. Rajpurohit and P. Jain (2017): SSR Based Molecular Characterization of Soybean [Glycine max (L.) Merrill] Genotypes. Bull. Env. Pharmacol. Life Sci., 7 (1): 46-54.

Johnson, H.W.; H.F. Robinson and R.E. Comstock (1955): estimates of genetic and environmental variability in soybean. Agron. J., 47: 314-318.

Karnwal, M.K. and K. Singh (2009): Studies of genetic variability, character association and path coefficient for seed yield and its contributing traits in soybean [Glycine max (L.)

Merrill]. Legume research, 32(1):70-73.

Khaled, A.G.A.; G.A.A. El-Sherbeny and H.S. Abdelghny (2019a): SRAP and ISSR molecular markers-trait associations in sorghum genotypes. Assiut J. Agric. Sc., 50(2): 72-73.

Khaled, A.G.A.; G.A.R. El-Sherbeny and H.M.A. Elsayed (2019b): Three SRAP Molecular Markers Linked to Yield Component Traits in Wheat. Asian Journal of Research and Review in Agriculture, 1(1): 25-35.

Khaled, A.G.A.; K. Hamam (2015): Association of molecular markers with phenotypic traits of bread wheat genotypes. Egypt. J. Genet. Cytol.,44:115-130.

Mahbub M.M. ; M. M. Rahman; M. S. Hossain; F. Mahmud and M.M. Kabir (2015): Genetic Variability, Correlation and Path Analysis for Yield and Yield Components in Soybean. American-Eurasian Journal of Agricultural and Environment Science, 15 (2): 231-236.

Mantel, N. (1967): The detection of disease clustering and a generalized regression approach. Cancer Res., 27:209-220.

Meng, H.; S. Chen; Z. Cheng; D. Chai and Y. LI (2012): SRAP Markers for Fruit Shape in Cucumber. Pak. J. Bot., 44(4): 1381-1384.

Ntuli, N.R.; P.B. Tongoona and A.M. Zobolo (2015): Genetic diversity in Cucurbita pepo landraces revealed by RAPD and SSR markers. Scientia Horticulturae, 189:192-200.

Pandey, K.; K. Singh; B.V. Singh; P. Bhartiya; M.K. Gupta and N.S. Yadav (2008): Character association and path coefficient analysis in advance breeding lines of soybean [Glycine max (L.) Merrill.]. Soybean Research, 45:34-38.

Pathan, S.M.; J.D. Lee; D.A. Sleper; F.B. Fritschi; R.E. Sharp; T.E. Carter; R. L. Nelson; C.A. King; W.T. Schapaugh; M. R. Ellersieck; H.T. Nguyen and J.G. Shannon (2014): Two soybean plant introductions display slow leaf wilting and reduced yield loss

under drought. J. Agron. Crop Sci., 200:231–236.

Peng, J.H. and N.L.V. Lapitan (2005): Characterization of EST-derived microsatellites in the wheat genome development of ESSR markers. Funct. Integr. Genomics,5:80-96.

Poresbski, S. L.; G. Bailey and R. B. Baum (1997): Modification of CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. Plant Mol. Biol. Reporter, 12:8-15.

Powell, W.; M. Morgante; C. Andre; M. Hanafey; J. Vogel; S. Tingey and A. Rafalski (1996): The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. Molecular Breeding, 2: 225-238.

Robarts, D.W.H. and A.D. Wolfe (2014): Sequence-related amplifyed polymorphism (SRAP) markers. Applications in Plant Sciences,2 (7): 1400017.

Rohlf, F.J. (2000): NTSYS-pc: Numerical taxonomy and multivariate analysis system. Version 2.1 Exeter Software, Setauket, USA.

Sambrook, J.; E.F. Fritsch and T. Maniatis (1989): Molecular cloning: a laboratory manual. 2nd ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Sensoy, S.; S. Buyukalaca and K. Abak (2007): Evaluation of genetic diversity in Turkish melons (Cucumis melo L.) based on phenotypic characters and RAPD markers. Genetic Resources Crop Evolution, 54 (6): 1351-1365.

Snedecor, G.W. and W.G. Cochran (1967): Statistical methods. The IOWA State Univ. Press Ames. IOWA, USA pp : 557-564.

Steel, R.G. and J.H. Torrie (1980): Principal and Procedures of Statistics. Mc Grow Hill Book Inc., New York, USA.

Sun, B.; C. Fu; C. Yang; Q. Ma; D. Pan and H. Nian (2013): Genetic Diversity of Wild Soybeans from Some Regions of Southern China Based on SSR and SRAP Markers. American Journal of Plant Sciences, 4: 257-268.

Tatikonda, L.; S.P. Wani; S . Kannan; N. Beerelli; T.K. Sreedevi; D.A. Hoisington; P.

Devi and R.A.Varshney (2009): AFLP-based molecular characterization of an elite germplasm collection of Jatropha curcas L.: A biofuel plant. Plant Sci., 176:505-513.

Thudi, M.; R . Manthena; S.P. Wani; L . Tatikonda; D.A. Hoisington and R.A. Varshney (2010): Analysis of genetic diversity in Pongamia (Pongamia pinnata L. Pierre) using AFLP markers. J. Plant Biochem. Biot, 19:209-216.

Wang, M.; W. Du; S. Fu and D. Yu (2009): Mapping QTLs for seed yield and drought susceptibility index in soybean (Glycine max L.) across different environments. J. Genet. Genomics,36:721–731.

Yadav, N.S. (2005): Estimation of genetic diversity and characters association for yield and its components in elite lines of soybean [Glycine max (L.) Merril]. Unpublished M.Sc. Thesis, pp: 1-124.

Yıldız, M.; E. Ekbiç; E. Düzyaman; S. Serçe and K. Abak (2016): Genetic and phenotypic variation of Turkish Okra (*Abelmoschusesculentus* L. Moench) accessions and their possible relationship with American, Indian and African germplasms. J. Plant Biochem. Biotechnol., 25(3):234–244.