Cell Biology
2023; 11(2): 20-29
http://www.sciencepublishinggroup.com/j/cb
doi: 10.11648/j.cb.20231102.12
ISSN: 2330-0175 (Print); ISSN: 2330-0183 (Online)

Genetic Variation of Recombinant Inbred Lines Soybean (*Glycine max* L. Merrill) from USA at Jimma, Ethiopia

Gada Gudina[1, *], Abush Tesfaye[2], Amsalu Nebiyu[3], Getachew Bekele[1]

[1]Department of Plant Science, College of Agriculture, Mettu University, Mattu, Ethiopia

[2]Ethiopian Institute of Agricultural Research, Jimma Agricultural Research Center, Jimma, Ethiopia

[3]Department of Horticulture and Plant Science, College of Agriculture and Veterinary medicine, Jimma University, Jimma, Ethiopia

Email address:

gudinagada@gmail.com (Gada Gudina)

[*]Corresponding author

To cite this article:

Gada Gudina, Abush Tesfaye, Amsalu Nebiyu, Getachew Bekele. (2023). Genetic Variation of Recombinant Inbred Lines Soybean (*Glycine max* L. Merrill) from USA at Jimma, Ethiopia. *Cell Biology*, *11*(2), 20-29. https://doi.org/10.11648/j.cb.20231102.12

Received: November 6, 2023; **Accepted**: November 21, 2023; **Published**: December 22, 2023

Abstract: In this study the genetic variability of soybean lines generated from segregating populations introduced from USA was evaluated. A total of 97 soybean genotypes that were introduced from USA along with three checks were grown in 10×10 simple lattice design with two replications at Jimma, Ethiopia. The ANOVA results showed significant ($p \leq 0.05$) variations in days to flowering, days to maturity, plant height, number of branches per plant, number of pods per plant, pod length, number of seeds per pod, number of seeds per plant, 100 seed weight, above ground biomass, harvest index, and grain yield indicating a considerable variability among the tested genotypes for the characters. Characters viz., plant height, number of branches per plant, above ground biomass and grain yield had high heritability and high genetic advance. Grain yield had positive and high significant ($p \leq 0.01$) genotypic correlations with harvest index (0.746) and 100 seed weight (0.267). Similarly, grain yield showed positive and significant ($p \leq 0.05$) genotypic associations with number of seeds per plant (0.225) and above ground biomass (0.205). This implies that higher mean values for these traits tend to improve grain yield in soybean. Cluster analysis grouped the genotypes into three clusters with the maximum squared distance found between cluster II and III. The principal component analysis revealed that the first four principal components (PCs) accounted for more than 71.25% of the total variation. The variability amongst the tested genotypes, heritability and genetic advance, as well as the associations in the tested traits provide information for an increased soybean productivity using this lines.

Keywords: Genetic Advance, Variability, Principal Components, RIL, Soybean

1. Introduction

Soybean (*Glycine max* L. Merrill) is among a worldwide cultivated legume crop that belongs to the Leguminosae family, in the subfamily, Papilionideae and genus, Glycine. It is originated from the Northeastern China, and distributed throughout Asia, USA, Brazil, and Argentina. It has 20 chromosome pairs (2n = 2x = 40) and is a self-pollinated species with less than a percent out-crossing [31]. Soybean has enormous importance as nutritionally rich human food and livestock feed, as well as other industrial and commercial uses. It is classified as an oilseed crop, containing significant amounts of all the essential amino acids, minerals, and vitamins for human nutrition. It has the highest protein content (40-42%) than any other grain crops, with oil content ranging 20-22% [15].

According to the United Nation Food and Agriculture Organization report (2017) the total world soybean production was 320.15 million metric tons. The three-leading soybean producing countries in the world are USA, Brazil and Argentina; while South Africa, Nigeria and Zambia are the top ranked soybeans producers in Africa. Soybean is introduced in Ethiopia the 1950's [6] and currently soybean production in the country is estimated to cover 38,072.7 ha of land with mean yield 2.27 t ha^{-1} at national level [11]. Soybean productivity in the country seems very low regarding the potential productivity provided by research evidence, which might reach up to 4.0 t ha^{-1} [34]. The main

reasons for the low productivity of the crop include: lack of high yielding, disease resistance and well adapted improved varieties, narrow genetic base, low yielding potential of the released varieties under production and meagre soil fertility [14].

Understanding the genetic variability of a crop is of great importance, in making selection, effective and easier, through enabling the selection of superior genotypes in diverse generations. Yield is a complex character controlled by several genes and affected by environment [4]. The variability in the yield characters is the consequence of genetic and environmental effect. Thus, it is necessary to partition the variability into heritable and non- heritable components measured as genotypic coefficient of variation (GCV), phenotypic coefficient of variation (PCV), heritability and genetic advance expressed as percent of mean.

For the selection for desirable attributes to be realized there should be sufficient genetic variability in the existing genetic materials. Therefore, generating the genetic variability and assessment of this variability is a vital step in crop improvement programs. Jimma Agricultural Research Center in Ethiopia introduced soybean segregating generations from the breeding program of University of Illinois in the year 2014 and have made plant to row selections in the year 2017. Of these, plant to row selections, some of the lines that best performed under Jimma conditions were used for this study. Thus, the study is designed to estimate the extent of genetic variability of the soybean lines that were generated from the segregating populations introduced from USA and evaluate their performance for further studies and breeding program. More specifically the study was aimed to evaluate the association of various characters with grain yield, and among themselves through correlation studies at both phenotypic and genotypic levels, to estimate the genetic divergence of the lines using D^2 analysis, cluster analysis and principal component analysis to investigate the main characters causative to the genetic variability.

2. Materials and Methods

2.1. Description of the Experimental Site

The field experiment was conducted at Jimma Agricultural Research Center [19], Jimma Zone; Oromia Regional State, Southwestern Ethiopia in the 2019 main cropping season. JARC is located about 365 km away from Addis Ababa, and situated at an altitude of 1754 m.a.s.l., latitude of 7° 40' N and longitude of 36° 47' E. The specific soil types of JARC are Upland: Chromic Nitosol and Cambisol; Bottom land: Fluvisol, with pH of 6.2 - 6.8 and the agro-ecology is characterized as sub humid type of climate. The annual average rain fall of the center is 1572 mm; while the minimum and maximum temperatures of the center are 11.6°C and 26.3°C, respectively [14].

2.2. Experimental Materials

A total of 97 lines generated from the segregating populations of the soybean breeding program of University of Illinois, USA, in the year 2014 as shown in (Table 1) were used as experimental materials. The materials are introduced at the stage of F3, and their segregation is fixed through successive selfing and selection of single plants starting from F4, using modified single seed descent breeding method at Jimma Agricultural Research Center.

Table 1. *Experimental materials.*

No	Genotypes	Pedigree		
1	T35-15 T77-16-SB1	LD07-3395bf	X	TGx 1987-10F
2	T50-15-T139-16-SA1	Benning	X	TGX 1987-129F
3	T75-15-T240-16-SA1	LD12-10534	X	TGX 1805-31F
4	T54-15-T156-16-SB1	G00-3213	X	TGX 1740-2F
5	T74-15-T239-16-SE1	LD07-3395bf	X	TGX 1987-14F
6	T34-15-T70-16-SA1	LD07-3395bf	X	TGX 1904-6F
7	T28-15-T61-16-SC1	LD11-7311	X	TGX 1989-19F
8	T28-15-T62-16-SD1	LD11-7311	X	TGX 1989-19F
9	T55-15-T159-16-SA1	NC-Raleigh	X	TGX 1740-2F
10	T46-15-T120-16-SB1	LD12-10534	X	TGX 1988-5F
11	T27-15-T53-16-SC1	LD11-7311	X	TGX 1987-14F
12	T51-15-T142-16-SA1	G00-3213	X	TGX 1987-129F
13	T27-15-T51-16-SA3	LD11-7311	X	TGX 1987-14F
14	T76-15-T241-16-SC1	LD12-10534	X	TGx 1987-10F
15	T57-15-T167-16-SB1	SC10-57	X	TGx 1987-10F
16	T42-15-T97-16-SA1	LD12-10534	X	TGX 1740-2F
17	T19-15-T38-16-SC1	LD12-00268	X	TGX 1485-1D
18	T50-15-T139-16-SB2	Benning	X	TGX 1987-129F
19	T27-15-T51-16-SA2	LD11-7311	X	TGX 1987-14F
20	T45-15-T116-16-SC2	LD12-10534	X	TGX 1987-76F
21	T54-15-T157-16-SC1	G00-3213	X	TGX 1740-2F
22	T35-15-T79-16-SD1	LD07-3395bf	X	TGx 1987-10F
23	T44-15-T111-16-SI2	LD12-10534	X	TGX 1835-10E
24	T50-15-T141-16-SC1	Benning	X	TGX 1987-129F
25	T74-15-T239-16-SE2	LD07-3395bf	X	TGX 1987-14F

No	Genotypes	Pedigree		
26	T44-15-T106-16-SD1	LD12-10534	X	TGX 1835-10E
27	T55-15-T160-16-SB2	NC-Raleigh	X	TGX 1740-2F
28	T44-15-T107-16-SE1	LD12-10534	X	TGX 1835-10E
29	T74-15-T240-16-SF1	LD07-3395bf	X	TGX 1987-14F
30	T44-15-T108-16-SF1	LD12-10534	X	TGX 1835-10E
31	T35-15-T80-16-SE1	LD07-3395bf	X	TGx 1987-10F
32	T74-15-T240-16-SF2	LD07-3395bf	X	TGX 1987-14F
33	T45-15-T116-16-SA1	LD12-10534	X	TGX 1987-76F
34	T27-15-T52-16-SB1	LD11-7311	X	TGX 1987-14F
35	T33-15-T68-16-SA1	LD07-3395bf	X	TGX 1440-1E
36	T47-15-T124-16-SC1	LD12-10534	X	TGX 1990-67F
37	T73-15-T230-16-SE1	LD07-3395bf	X	TGX1835-10E
38	T56-15-T164-16-SB1	N05-7432 IB	X	TGX 1740-2F
39	T27-15-T57-16-SG1	LD11-7311	X	TGX 1987-14F
40	T52-15-T147-16-SA1	Cook	X	TGx 1987-10F
41	T66-15-T193-16-SB1	Conquista	X	TGX 1740-2F
42	T15-15-T20-16-SH1	LD12-10950	X	TGX 1988-5F
43	T6-15-T2-16-SA2	LD12-07686	X	TGX 1440-1E
44	T29-15-T63-16-SA1	LD11-7311	X	TGX 1991-10F
45	T26-15-T50-16-SB1	LD11-7311	X	TGX 1987-76F
46	T54-15-T155-16-SA2	G00-3213	X	TGX 1740-2F
47	T56-15-T165-16-SC1	N05-7432 IB	X	TGX 1740-2F
48	T25-15-T47-16-SA1	LD11-7311	X	TGX 1835-10E
49	T25-15-T48-16-SB1	LD11-7311	X	TGX 1835-10E
50	T6-15-T223-16-SB	FT-Abyara	X	TGX 1835-10E
51	T16-15-T-31-16-SK1	LD12-10950	X	TGX 1990-67F
52	T49-15-T134-16-SC1	G00-3213	X	TGX 1740-2F
53	T37-15-T81-16-SA1	LD07-3395bf	X	TGX 1989-19F
54	T27-15-T51-16-SA1	LD11-7311	X	TGX 1987-14F
55	T24-15-T46-16-SA2	LD11-7311	X	TGX 1485-1D
56	T63-15-T183-16-SB1	Padre	X	TGX 1987-62F
57	T52-15-T147-16-SA2	Cook	X	TGx 1987-10F
58	T53-15-T153-16-SC1	Prichard	X	TGx 1987-10F
59	T23-15-T45-16-SA1	LD11-7311	X	TGX 1448-2E
60	T52-15-T148-16-SB1	Cook	X	TGx 1987-10F
61	T70-15-T219-16-SF2	FT-Abyara	X	TGX 1740-2F
62	T74-15-T236-16-SB1	LD07-3395bf	X	TGX 1987-14F
63	T47-15-T122-16-SA1	LD12-10534	X	TGX 1990-67F
64	T48-15-T128-16-SA1	Benning	X	TGX 1740-2F
65	T73-15-T228-16-SC1	LD07-3395bf	X	TGX 1835-10E
66	T73-15-T231-16-SF1	LD07-3395bf	X	TGX 1835-10E
67	T56-15-T163-16-SA1	N05-7432 IB	X	TGX 1740-2F
68	T70-15-T220-16-SG1	FT-Abyara	X	TGX 1740-2F
69	T44-15-T111-16-SI1	LD12-10534	X	TGX 1835-10E
70	T24-15-T46-16-SA1	LD11-7311	X	TGX 1485-1D
71	T67-15-T203-16-SG2	Conquista	X	TGX 1989-19F
72	T27-15-T58-16-SH2	LD11-7311	X	TGX 1987-14F
73	T16-15-T23-16-SC1	LD12-10950	X	TGX 1990-67F
74	T27-15-T58-16-SH1	LD11-7311	X	TGX 1987-14F
75	T19-15-T39-16-SD1	LD12-00268	X	TGX 1485-1D
76	T51-15-T146-16-SE1	G00-3213	X	TGX 1987-129F
77	T55-15-T161-16-SC3	NC-Raleigh	X	TGX 1740-2F
78	T47-15-T122-16-SA1	LD12-10534	X	TGX 1990-67F
79	T27-15-T57-16-SG2	LD11-7311	X	TGX 1987-14F
80	T71-15-T224-16-SC1	FT-Abyara	X	TGX 1835-10E
81	T29-15-T64-16-SB1	LD11-7311	X	TGX 1991-10F
82	T49-15-T132-16-SA1	G00-3213	X	TGX 1740-2F
83	T53-15-T152-16-SB1	Prichard	X	TGx 1987-10F
84	T46-15-T119-16-SA1	LD12-10534	X	TGX 1988-5F
85	T52-15-T149-16-SC1	Cook	X	TGx 1987-10F
86	T16-15-T27-16-SG3	LD12-10950	X	TGX 1990-67F
87	T39-15-T92-16-SC2	LD12-10534	X	TGX 1448-2E
88	T35-15-T78-16-SC1	LD07-3395bf	X	TGx 1987-10F
89	T71-15-T222-16-SA1	FT-Abyara	X	TGX 1835-10E
90	T15-15-T16-16-SD1	LD12-10950	X	TGX 1988-5F
91	T72-15-T225-16-SA1	LD11-7311	X	TGX 1740-2F
92	T33-15-T69-16-SB1	LD07-3395bf	X	TGX 1440-1E

No	Genotypes	Pedigree		
93	T39-15-T96-16-SG1	LD12-10534	X	TGX 1448-2E
94	T62-15-T181-16-SA1	Padre	X	TGx 1987-10F
95	T25-15-T47-16-SA2	LD11-7311	X	TGX 1835-10E
96	T54-15-T155-16-SA1	G00-3213	X	TGX 1740-2F
97	T16-15-T28-16-SH1	LD12-10950	X	TGX 1990-67F
98	Nyala (check 1)			
99	Afgat (check 2)			
100	Clarck 63 k (check 3)			

2.3. Experimental Design

The 100 entries were laid out in 10 × 10 simple lattice design with two replications. There were four rows in each plot of 4m × 2.4m with a total plot size of 9.6m^2. The inter-row spacing was 60 cm, while the between plants spacing was 5cm. A distance of 1m between adjacent plots and 1.5m between blocks were maintained for free movement and plot management. Planting was done in June 2019 under rainfed conditions. All the required management practices were applied the plots uniformly.

2.4. Data Collection

The central two rows of each plot's plants were used to calculate the crop phenology data (days to 50% flowering and days to physiological maturity) in accordance with the descriptors for soybean. A random sample of 10 plants from the two middle rows were chosen to represent each crop growth characteristic, such as plant height (cm) and the number of primary branches per plant, as well as the crop yield components, such as the number of pods per plant, pod length, number of seeds per pod, and number of seeds per plant. From each plot's seed harvest, 100 seed were taken, and their weight was assessed in grammes using a sensitive balance. At harvest, the two middle rows were used to measure the yield characteristics (above ground biomass at constant weight and grain yield at a standard moisture content of 12.5%). Grain weight divided by above ground biomass production multiplied by 100 was used to construct harvest index.

2.5. Statistical Analysis

The existence of statistically significant differences among genotypes for the parameters under consideration was tested by analysis of variance (ANOVA), using SAS version 9.3 (SAS institute, 2012). Least Significant Difference (LSD) at $p \leq 0.05$ was employed for mean separation. The genotypic parameters and heritability estimation and genetic advance were calculated according to the formula suggested by Burton and De Vane (1953). Analysis of phenotypic and genotypic correlation coefficients according to the formula proposed by [20, 33], and Clustering was carried out using the Proc cluster procedure of SAS version 9.3 [30] and the average linkage clustering technique of the observations were employed.

3. Results and Discussions

3.1. Analysis of Variance

The ANOVA result revealed that the genotypes showed a highly significant difference ($p \leq 0.01$) for days to 50% flowering, days to maturity, plant height, numbers of branches, pods, and seeds per plant, above ground biomass, harvest index, hundred seed weight and grain yield. In addition the difference among genotypes was significant ($p \leq 0.05$) for traits, such as, pod length and number of seeds per pod indicating considerable genetic variability among the tested genotypes for these characters (Table 2).

In agreement with this finding, [27] observed a highly significant genetic variation for plant height, number of branches per plant and 100 seed weight; [1] for days to 50% flowering, days to maturity and number of pods per plant in soybean. [18] also found a significant ($p \leq 0.05$) genetic variation for pod length and number of seeds per pod in soybean.

3.2. Range and Mean of the Characters

The range and mean performance of the studied characters are presented in Table 3. Wide ranges were recorded for days to 50% flowering (61-91.5), days to maturity (123.5-168), plant height (53.1-105.9cm), numbers of branches (1.9-6.4), pods (19.9-56.6), and seeds (39.4-104.1) per plant, above ground biomass (3050.5-5508.2gm plot^{-1}), harvest index (9.9-43.2), 100 seed weight (10.4-20.2g) and grain yield (1114.6-3869.7kg ha^{-1}). This implies the wide differences in the experimental materials for these characters (Table 3). On the other hand, the three released varieties (standard check varieties) used in the present study such as Nyala (2020.2 kg ha^{-1}), Afgat (2337.3 kg ha^{-1}), and Clark 63 K (2321.5 kg ha^{-1}) produced less than the grand mean for grain yield (2388.5 kg ha^{-1}).

On the basis of the average grain yield of the genotypes, T27-15-T53-16-SC1 (3869.7 kg ha^{-1}), T27-15-T52-16-SB1 (3577.6 kg ha^{-1}), T74-15-T240-16-SF1 (3473.1 kg ha^{-1}), T27-15-T58-16-SH2 (3323.7 kg ha^{-1}) and T19-15-T38-16-SC1 (3314.9 kg ha^{-1}) are high yielding. These genotypes could be advanced in variety trials.

Table 2. *Mean squares for the different sources of variation and the corresponding CV (%) for the characters studied on genetic variation.*

Traits	Replication	Block (Rep.)	Genotypes (adj.)	Genotypes (un adj.)	Error	C.V.%	RCBD error
Days to 50% flowering	60.5	7.2	91.98**	98.8	4.04	2.78	4.62
Days to maturity	330.2	19.8	242.68**	276.11	19.7	2.94	19.74
Plant height (cm)	341.4	17.6	176.2**	211.13	26.37	6.9	24.78
Number of branches per plant	0.001	0.16	1.56**	1.75	0.21	11.4	0.22
Number of pods per plant	689.3	23.96	82.38**	85.42	30.16	15.8	29.04
Pod length (cm)	0.44	0.22	0.19*	0.2	0.12	10.6	0.14
Number of seeds per plant	1420.4	78.69	251.3**	270.82	79.32	14	79.21
Number of seeds per pod	0.35	0.024	0.023*	0.025	0.016	5.44	0.017
Above ground biomass (gm plot^{-1})	11414498	207355.3	634507**	691267	121576	8.15	137173
Harvest index	403.9	10.92	66.65**	73.3	18.93	16.7	17.48
100 seed weight (gm)	11.18	2.64	5.21**	5.5	2.08	9.82	2.19
Grain yield (kg ha^{-1})	9612889	91145.3	430820.6**	479772	91608.8	12.67	91625

**= highly significant, *= significant

Table 3. *Estimates of mean, range, variance components, broad sense heritability, genotypic and phenotypic coefficient of variability, genetic advance as percent of mean for the 12 characters of soybean genotypes studied for genetic variation.*

Traits	Range	Mean	σ^2g	σ^2e	σ^2p	GCV (%)	PCV (%)	H^2 (%)	GA K=5%	GAM K=5%
DF	61-91.5	72.19	43.97	4.04	48.01	9.2	9.6	91.58	1309.35	18.14
DM	123.5-168	150.9	111.49	19.7	131.19	7	7.6	84.98	2007.43	13.3
PH	53.1-105.9	74.37	74.91	26.37	101.28	11.64	13.53	73.96	1535.5	20.64
BR	1.9-6.4	4.1	0.68	0.21	0.885	20.11	22.9	76.27	147.91	36.07
POD	19.9-56.6	34.76	26.11	30.16	56.27	14.7	21.6	46.4	718	20.65
PDL	2.8-4.44	3.34	0.16	0.125	0.285	12	16	56.14	61.4	18.38
SPL	39.4-104.1	63.6	85.99	79.32	165.31	14.6	20.22	52.08	1378.96	21.68
SPO	2.1-2.7	2.31	0.0035	0.016	0.0195	2.56	6.05	18	5.16	2.23
AGB	3050.5-5508.2	4275.17	256466	121576	378042	11.85	14.38	67.8	86051.4	20.13
HI	9.9-43.2	26.02	23.86	18.93	42.79	18.77	25.14	55.76	752.32	28.9
HSW	10.4-20.2	14.69	1.57	2.08	3.65	8.53	13	42.9	169.18	11.51
YLD	1114.6-3869.7	2388.5	169606	91609	261215	17.24	21.4	64.9	68460.5	28.66

DF=Days to 50% flowering, DM=Days to maturity, PH= Plant height, BR= Number of branches per plant, POD= Number of pods per plant, PDL= Pod length, SPL= Number of seeds per plant, SPO= Number of seeds per pod, AGB= Above ground biomass, HI= Harvest index, HSW= 100 seed weight, YLD= Grain yield, σ^2g = Genotypic variance, σ^2e = Environmental variance, σ^2p = Phenotypic variance, h^2 = Broad sense heritability, GCV= Genotypic coefficient of variability, PCV= Phenotypic coefficient of variability, GA= Genetic Advance, GAM= Genetic advance as percent of mean and K = Selection intensity

3.3. Phenotypic and Genotypic Variation

In the estimate of variance components, the phenotypic coefficient of variance (PCV) and genotypic coefficient of variance (GCV) of the characters studied are presented in Table 3. The PCV varied from 6.05% to 25.14% and GCV from 2.56% to 20.11%. In general, PCV and GCV values were classified as low, medium and high with respective values of 0-10%, 10-20% and >20% in soybean [20].

Accordingly, high PCV values were recorded for number of branches per plant, number of pods and number of seeds per plant, harvest index and grain yield. High GCV value was recorded only for number of branches per plant. Moderate PCV values were recorded for plant height, pod length, above ground biomass and 100 seed weight, while the GCV values were moderate for characters, such as, plant height, number of pods per plant, pod length, number of seeds per plant, above ground biomass, harvest index and grain yield. Low PCV values were observed for days to 50% flowering, days to maturity and number of seeds per pod, while the GCV values were for traits, such as, days to 50% flowering, days to maturity, 100 seed weight and number of seeds per pod.

In agreement with the present study, [4] reported high PCV values on soybean for number of pods and number of seeds per plant, and moderate GCV value for number of seeds per plant.

The genetic variance was found relatively greater than its corresponding environmental variance for days to 50% flowering, days to maturity, plant height, number of branches per plant, pod length, number of seeds per plant, above ground biomass, harvest index and grain yield. This means that in the phenotypic expression of these traits, the effect of environmental factors was low. However, the magnitudes of genotypic variances were smaller than that of environmental variance for number of pods per plant, number of seeds per pod and 100 seed weight. This shows high environmental effect on the phenotypic expression of these traits.

3.4. Heritability and Genetic Advance

The broad sense heritability for the characteristics under investigation in this study was computed and is shown in Table 3. The range was 18% for the number of seeds per pod, 91.58% for the number of days, and 50% for flowering. According to [20], heritability values can be classified as low (0-30%), moderate (30-60%), or high (60% and beyond). Days to 50% flowering, days to maturity, plant height, number of branches per plant, above-ground biomass, and

grain yield all showed high heritability estimates based on these criteria. This indicates that the genotypic component of these traits' inheritance predominates, giving breeders the chance to respond to any intense selection and making phenotypic performance-based selection relatively simple, improving soybean for these traits. While seeds per pod showed low heredity estimates, the number of pods and seeds per plant, pod length, harvest index, and 100 seed weight showed medium heritability estimates, showing a substantial influence of environmental factors. In agreement with the present study, [29] stated that the highest heritability was recorded for days to flowering, days to maturity and plant height. Furthermore, [27] found moderate heritability estimates for seeds per plant in soybean. [4] also reported high heritability for number of branches per plant, moderate heritability for pods per plant, and low heritability for seeds per pod which is similar with the present study.

The genetic advance as percent of mean (GAM) at 5% selection intensity is also shown in Table 3. Estimates of GAM at 5% selection intensity ranged from 2.23% for number of seeds per pod to 36.07% for number of branches per plant. The GAM classified as low (0 to 10%), moderate (10 to 20%) and high (20% and above) as stated by [20]. Accordingly, high GAM was recorded for plant height, number of branches per plant, number of pods per plant, number of seeds per plant, above ground biomass, harvest index and grain yield. While days to 50% flowering, days to maturity, pod length and 100 seed weight observed moderate genetic advance as percent of mean, whereas number of seeds per pod showed low genetic advance as percent of mean.

High heritability for a given quantitative trait does not necessarily indicate high genetic advance. [2] stated that genetic advances together with heritability estimates would be more fruitful than heredity alone in predicting the outcome of selection. As a result, traits like plant height, the number of branches per plant, above-ground biomass, and grain yield were found to have high heritability along with high genetic advance; this indicates the presence of additive gene action for the expression of these traits, which is fixable in the next generation, and selection in the next population based on these characters would be ideal. Similarly, [15] reported traits combining such high heritability and genetic advance are predominantly controlled by additive gene action and can easily be improved by selection. [2] also found high heritability and high genetic advance for plant height and number of branches per plant.

3.5. Association of Characters

The results of Estimates of genotypic and phenotypic correlation coefficients between each pair of characters revealed that, in most cases, the genotypic correlation coefficients were higher than the phenotypic correlation coefficients (Table 4 and 5), which indicated the inherent association among various characters independent of environmental influence.

Grain yield had positive and significant ($p \leq 0.01$) genotypic correlations with harvest index (0.746) and 100 seed weight (0.267) (Table 4). Likewise, grain yield showed positive and significant ($p \leq 0.05$) genotypic correlations with number of seeds per plant (0.225) and above ground biomass (0.205). The current findings infer the selection of soybean genotypes with high values of 100 seed weight, harvest index, above ground biomass and number of seeds per plant incline to improve grain yield. Therefore, these characters could be taken into account as a selection in the soybean improvement program. Similar to the current study, [27] found a positive association between soybean seed yield and the number of pods, seeds per plant, and 100-seed weight. Similar findings were made by [18], who also found positive and highly significant genotypic relationships between soybean harvest index and grain yield. On the other hand, significant and negative genotypic correlations were detected between days to 50% flowering and grain yield (Table 4). This implies the inverse relationship between the mean values of this character and grain yield. Similarly, [5] reported negative and significant genotypic associations of grain yield and days to 50% flowering in soybean, which is similar with the present study.

At phenotypic level, number of pods per plant (0.272), number of seeds per plant (0.308), above ground biomass (0.279) and harvest index (0.736) had positive and significant ($p \leq 0.01$) correlations with grain yield (Table 5). However, number of seeds per pod (0.158) and 100 seed weight (0.125) had positive and significant ($p \leq 0.05$) phenotypic associations with grain yield. In line with the current study results, [27] testified highly significant positive phenotypic association of number of pods per plant and number of seeds per plant with grain yield in soybean.

Based on genotypic and phenotypic correlation coefficients (Tables 4 and 5), grain yield had non-significant but positive correlation with branches per plant. Furthermore, grain yield had non-significant but positive genotypic correlation with number of pods per plant and number of seeds per pod. However, branches number per plant had significant and positive genotypic correlations with seed number per plant and harvest index which is positive and significantly correlated with grain yield. In addition, pod number per plant and seed number of per pod had significant and positive phenotypic and genotypic correlation with seed number per plant which is also positive and significantly correlated with grain yield. Therefore, branches number per plant, pod number per plant and seed number per pod were indirectly used for improving grain yield.

On the basis of genotypic and phenotypic correlation coefficients, grain yield also showed non-significant and negative relationships with plant height and days to maturity. Additionally, non-significant, positive phenotypic, and negative genotypic correlations between grain yield and pod length were found (Tables 4 and 5). Grain yield and days to 50% flowering were shown to be negatively and non-significantly correlated at the phenotypic level (Table 5). [23] observed that days to maturity and plant height were inversely connected with grain yield in soybean, which is

similar to the findings of the current experiment.

Table 4. *Genotypic correlation coefficients for 12 characters studied for genetic variation.*

Variable	DM	PH	BR	POD	PDL	SPL	SPO	AGB	HI	HSW	YLD
DF	0.492**	0.403**	0.522**	0.043	-0.100	-0.037	-0.079	0.342**	-0.373**	-0.259**	-0.203*
DM		0.624**	0.404**	-0.164	-0.177	-0.312**	-0.223*	0.568**	-0.495**	0.005	-0.146
PH			0.310**	0.098	-0.018	-0.038	-0.093	0.437**	-0.419**	-0.091	-0.161
BR				0.343**	0.067	0.219*	0.071	0.204*	0.197*	-0.310**	0.100
POD					0.232*	0.840**	0.280**	-0.103	0.188	-0.265**	0.172
PDL						0.344**	0.867**	-0.098	0.027	-0.008	-0.020
SPL							0.394**	-0.196*	0.295**	-0.284**	0.225*
SPO								-0.064	0.062	-0.017	0.050
AGB									-0.535**	0.170	0.205*
HI										0.110	0.746**
HSW											0.267**

Table 5. *Phenotypic correlation coefficients for 12 characters studied for genetic variation.*

Variable	DM	PH	BR	POD	PDL	SPL	SPO	AGB	HI	HSW	YLD
DF	0.458**	0.38**	0.494**	0.07	-0.077	0.0006	-0.047	0.324**	-0.284**	-0.235**	-0.112
DM		0.59**	0.359**	-0.123	-0.147*	-0.261**	-0.144*	0.509**	-0.415**	-0.01	-0.085
PH			0.279**	0.142*	-0.025	0.027	-0.036	0.398**	-0.278**	-0.111	-0.05
BR				0.335**	0.107	0.228**	0.105	0.176*	0.149*	-0.233**	0.052
POD					0.255**	0.843**	0.316**	0.032	0.253**	-0.212**	0.272**
PDL						0.313**	0.824**	0.002	0.052	0.056	0.049
SPL							0.372**	-0.045	0.330**	-0.233**	0.308**
SPO								0.068	0.122	0.038	0.158*
AGB									-0.348**	0.085	0.279**
HI										0.046	0.736**
HSW											0.125*

*, ** indicates significance at 0.05 and 0.01 probability levels, respectively. DF=Days to 50% flowering, DM=Days to maturity, PH= Plant height, BR= Number of branches per plant, POD= Number of pods per plant, PDL=Pod length, SPL= Number of seeds per plant, SPO= Number of seeds per pod, AGB= Above ground biomass, HI= Harvest index, HSW= 100 seedweight, YLD= Grain yield

3.6. Cluster Analysis

The results of the cluster analysis showed that the three released varieties and 97 soybean genotypes could be categorised into three clusters. The largest cluster, cluster I, included 60 genotypes. Cluster II, which included 36 genotypes, was close behind, and cluster III, which included four genotypes (Table 6). The grouping pattern at the cluster level was mostly based on the comparable nature of the characteristics, regardless of the germplasm source. The standardized Mahalanobis statistics indicates the presence of genetic distance amongst the three clusters and showed significant difference ($P \leq 0.01$), except cluster I and II which is non-significant (Table 7). The longest cluster distance was found between cluster II and III, and cluster I and III; while the smallest cluster distance was found between cluster I and II (Table 7).

The absence or little genetic diversity between cluster I and II, shown by non-significant and minimum inter-cluster distance. This suggests that crossing the genotypes from the two clusters less contribute to higher heterotic value in F1 and thus the range of variability in the segregating F2 population becomes narrow.. However, the parents selected from divergent cluster groups expected to result in the maximum genetic recombination.. In the present study, crosses involving parents selected from cluster II and III will result in maximum recombination and segregation of progenies.

The cluster mean analysis showed that cluster I had the highest cluster mean for 100 weight, and produced the lowest cluster mean for number of pods per plant and number of seeds per plant (Table 8). It produced the second highest cluster mean for number of branches per plant and seeds per pod, pod length, seeds per pod and grain yield, while the third highest cluster mean for days to 50% flowering, days to maturity, plant height, above ground biomass and harvest index.

Cluster II, produced the highest cluster mean for pod length; the second highest cluster mean for number of pods per plant and harvest index; the third highest cluster mean for number of seeds per plant and seeds per pod and grain yield; and on contrary produced the lowest cluster mean for days to 50% flowering, days to maturity, plant height, number of branches per plant, 100 seed weight and above ground biomass.

Cluster III, was characterized by the highest cluster mean for numbers of branches and number of pods per plant, number of seeds per pod, above ground biomass, harvest index and grain yield; the second highest cluster mean for days to 50% flowering, days to maturity, plant height, number of seeds per plant and harvest index; and the third highest cluster mean for pod length.

In general, cluster mean analysis revealed that the genotypes in cluster I could be selected for crossing, with the

intention to develop hybrid with high value of 100 seed weight. Accordingly, genotypes in cluster II could be preferred for the purpose of crossing, when the intention is to develop pod length, early maturing and short genotypes, and cluster III for numbers of branches and pods per plant, number of seeds per pod, above ground biomass, harvest index and grain yield.

Table 6. *Distribution of 100 soybean genotypes in three cluster groups tested for genetic variation.*

Clusters	Number of genotypes	Proportion (%)	Name of genotypes
Cluster I	60	60	T35-15-T77-16-SB1, T27-15-T58-16-SH1, T27-15-T51-16-SA1, T71-15-T222-16-SA1, T26-15-T50-16-SB1, T25-15-T47-16-SA1, T42-15-T97-16-SA1, T27-15-T57-16-SG1, T16-15-T-31-16-SK1, T46-15-T119-16-SA1, T50-15-T139-16-SA1, T52-15-T147-16-SA1, T35-15-T80-16-SE1, T29-15-T63-16-SA1, T44-15-T108-16-SF1, T53-15-T153-16-SC1, T74-15-T240-16-SF2, T45-15-T116-16-SC2, T51-15-T142-16-SA1, T70-15-T220-16-SG1, T34-15-T70-16-SA1, T33-15-T68-16-SA1, T51-15-T146-16-SE1, T15-15-T16-16-SD1, T39-15-T96-16-SG1, T24-15-T46-16-SA2, T48-15-T128-16-SA1, T28-15-T61-16-SC1, T50-15-T139-16-SB2, T54-15-T157-16-SC1, T73-15-T230-16-SE1, T24-15-T46-16-SA1, T44-15-T107-16-SE1, T55-15-T160-16-SB2, T66-15-T193-16-SB1, T6-15-T223-16-SB, T72-15-T225-16-SA1, T23-15-T45-16-SA1, T47-15-T122-16-SA1, T16-15-T23-16-SC1, T74-15-T239-16-SE1, T49-15-T132-16-SA1, T56-15-T164-16-SB1, T56-15-T165-16-SC1, T33-15-T69-16-SB1, T45-15-T116-16-SA1, T35-15-T79-16-SD1, T25-15-T47-16-SA2, T46-15-T120-16-SB1, T71-15-T224-16-SC1, T74-15-T236-16-SB1, T63-15-T183-16-SB1, T16-15-T28-16-SH1, T54-15-T156-16-SB1, T28-15-T62-16-SD1, T47-15-T122-16-SA1, T52-15-T149-16-SC1, T19-15-T38-16-SC1, T29-15-T64-16-SB1, T74-15-T240-16-SF1, T54-15-T155-16-SA1
Cluster II	36	36	T16-15-T27-16-SG3, Afgat, T54-15-T155-16-SA2, T52-15-T148-16-SB1, T73-15-T231-16-SF1, T44-15-T111-16-SI2, T70-15-T219-16-SF2, T27-15-T51-16-SA2, T6-15-T2-16-SA2, T55-15-T159-16-SA1, T19-15-T39-16-SD1, T73-15-T228-16-SC1, Clarck 63 k, T15-15-T20-16-SH1, T35-15-T78-16-SC1, T75-15-T240-16-SA1, T56-15-T163-16-SA1, T74-15-T239-16-SE2, T47-15-T124-16-SC1, T27-15-T51-16-SA3, T52-15-T147-16-SA2, T76-15-T241-16-SC1, T55-15-T161-16-SC3, T25-15-T48-16-SB1, T67-15-T203-16-SG2, T53-15-T152-16-SB1, T50-15-T141-16-SC1, T39-15-T92-16-SC2, T44-15-T106-16-SD1, T62-15-T181-16-SA1, T49-15-T134-16-SC1, T37-15-T81-16-SA1, Nyala, T44-15-T111-16-SI1, T57-15-T167-16-SB1
Cluster III	4	4	T27-15-T53-16-SC1, T27-15-T52-16-SB1, T27-15-T58-16-SH2, T27-15-T57-16-SG2

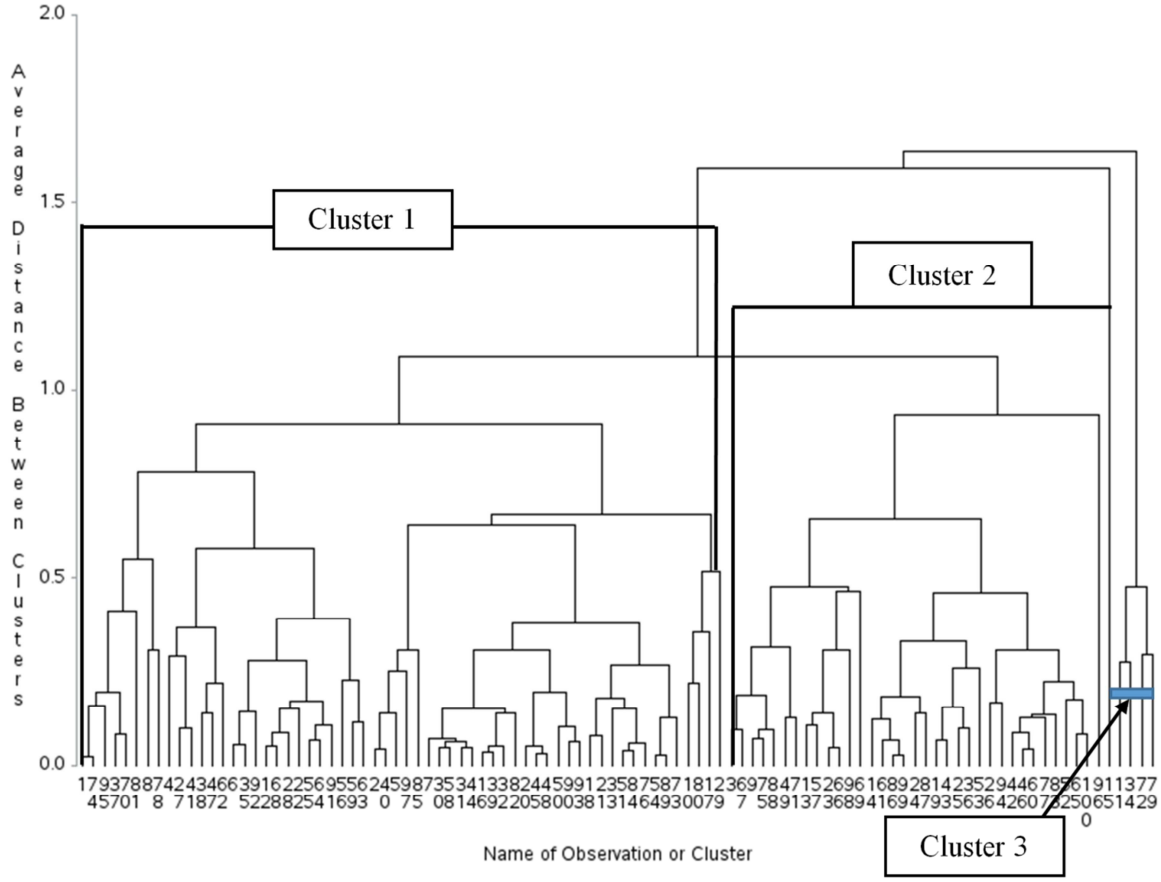

Figure 1. *Dendrogram of 100 soybean genotypes evaluated for genetic variation during 2019 main cropping season using 12 traits.*

Table 7. *Generalized squared distance between clusters 1-3 in 12 quantitative characters of soybean genotypes studied for genetic variation.*

Clusters	Cluster I	Cluster II	Cluster III
Cluster I	0	12.32ns	24.85**
Cluster II		0	66.54**
Cluster III			0

ns = non-significant, ** = significant at (P≤0.01)

Table 8. *Cluster mean for 12 quantitative characters of 100 soybean genotypes studied for genetic variation.*

Traits	Cluster I	Cluster II	Cluster III
DF	73.53	69.26*	75.18
DM	154.69	143.01*	162.13
PH	76.68	68.98*	83.54
BR	4.20	3.88*	4.60**
POD	34.02*	35.61	38.43**
PDL	3.34	3.35**	3.29
SPL	61.60*	66.26	68.88
SPO	2.32	2.31	2.35**
AGB	4559.53	3638.88*	5346.58**
HI	24.47	28.71	29.94**
HSW	15.02**	14.13*	14.95
YLD	2416.59	2250.46	3494.21**

* the lowest cluster mean, ** the highest cluster mean, DF=days to 50% flowering, DM=days to maturity, PH= Plant height (cm), BR= Number of branches per plant, POD= Number of pods per plant, PDL=Pod length (cm), SPL= Number of seeds per plant, SPO= Number of seeds per pod, AGB= Above ground biomass (gm $plot^{-1}$), HI= Harvest index, HSW= 100 seed weight (gm), YLD= Grain yield (kg ha^{-1})

3.7. Principal Component Analysis

Table 9. *Principal component score values, Eigen values and percent of variation explained by the first four PCs of soybean genotypes evaluated for genetic variation.*

Traits	PC 1	PC 2	PC 3	PC 4
Days to 50% flowering	0.360	0.214	0.046	-0.228
Days to maturity	0.448	0.007	0.257	-0.035
Plant height	0.373	0.152	0.141	0.175
Number of branches per plant	0.244	0.401	0.071	-0.165
Number of pods per plant	-0.126	0.550	0.043	0.135
Pod length	-0.053	0.032	0.375	-0.517
Number of seeds per plant	-0.215	0.531	0.013	0.135
Number of seeds per pod	-0.136	0.259	-0.077	0.491
Above ground biomass	0.352	-0.019	0.380	0.304
Harvest index	-0.439	0.042	0.293	-0.201
100 seed weight	-0.066	-0.338	0.344	0.459
Grain yield	-0.259	0.043	0.640	0.026
Eigen value	3.4	2.43	1.58	1.14
Variance explained (%)	28.4	20.24	13.13	9.48
Cumulative variance explained (%)	28.4	48.63	61.76	71.24

In the present study, PCA was performed for economic traits of soybean as shown in Table 9. Among the twelve, only four principal components (PCs) showed more than 1.00 Eigen value, which all together contributed to about 71.25% of the total variation in the studied population. The first PC that explained more than 28.4% of the total variation was influenced by the average values of the PC scores of days to 50% flowering, days to maturity, plant height, above ground biomass and harvest index. Given the higher percentage of variance provided by the first PCs to the overall variation, this suggests that these qualities are the main causes of the total variation in the genotypes under study.

Traits, such as the number of branches, pods, and seeds per plant, influenced the second PC, which contributed to 20.24% of the overall variation. The average values of the PC score for pod length, above ground biomass, 100 seed weight, and grain yield had an impact on the third PC, which contributed to 13.13 percent of the overall variation. The average values of the PC scores for pod length, number of seeds per pod, above ground biomass, and 100 seed weight had an impact on the fourth PC, which accounted for 9.48% of the overall variation (Table 9).

4. Conclusion

The overall present study results indicated the existence of considerable variability among the tested genotypes for these characters. The genetic advance, heritability, and correlations in the evaluated variables of the genotypes all supported the prospect of boosting soybean output. Therefore, for further soybean yield improvement programmes, selection and hybridization on these genotypes based on variables with moderate to high GCV, high heritability and genetic progress, and strong positive correlation coefficient can be advised. It is advised that future research into the variation of these genotypes for qualitative aspects like protein and oil content would be important.

Conflicts of Interest

The authors declare there is no conflict of interest.

References

[1] Agdew, B. and Getnet, A., 2006. Desirable Traits Influencing Grain Yield in Soybean. Innovative Systems Design and Engineering. (www.iiste.org). Accessed on February 10, 2013.

[2] Agrawal, A. P., Patil, S. A. and Math, P. M. S., 2001. Variability, heritability and genetic advance of some quantitative characters over the seasons in soybean. *Madras Agricultural Journal*, *88* (1/3): 36-39.

[3] Allard, R. W., 1960. Principles of Plant Breeding. John Wiley and Sons. Inc. New York. Genetic Variability. (http://en.wikipedia.org/wiki/). Accessed on November 13, 2011.

[4] Amarnath, K. C. N., Viswanatha, S. R. and Shivashankar, G., 1991. Genotypic and phenotypic variability and heritability of some quantitative characters in soybean (Glycine max (L.) Merill). *Mysore J. Agric. Sci*, *25*: 26-31.

[5] Arshad, M., Ali, N. and Ghafoor, A., 2006. Character correlation and path coefficient in soybean Glycine max (L.) Merrill. *Pakistan Journal of Botany*, *38* (1): 121.

[6] Asfaw, A., A. Tesfaye, S. Alamrie and M. Atnaf., 2006. Food and Forage Legumes of Ethiopia: progress and prospects. Proceedings of the workshop on food and forage legume. Soybean genetic improvement in Ethiopia, pp. 22-26.

[7] Bhatt, G. M., 1973. Comparison of various methods of selecting parents for hybridization in common bread wheat (Triticum aestivum L.). *Australian Journal of Agricultural Research*, *24* (4): 457-464.

[8] Burton, G. W. and Devane, E. H., 1953. Estimating heritability in tall fescue (Festuca arundinacea) from replicated clonal material 1. *Agronomy Journal*, *45* (10): 478-481.

[9] Cooper, M. C. and Milligan, G. W., 1988. The effect of measurement error on determining the number of clusters in cluster analysis. In *Data, expert knowledge and decisions* (pp. 319-328). Springer, Berlin, Heidelberg.

[10] Crossa, J., 1995. The use of multivariate methods in developing a core collection. *Core collections of plant genetic resources* Pp. 77-92.

[11] CSA (Central Statistical Authority). 2017. Agricultural sample survey statistical volume 1. Bulletin no 586. Addis Ababa, Ethiopia. P 14.

[12] Falconer, D. S. and Mackay, T. F. C., 1996. Introduction to quantitative genetics. Longman. *Essex, England*.

[13] FAOSTAT (Statistical database of the food and agriculture of the united nations) 2016. Statistics Division online database. www.fao.org/faostat (Assessed on May 2017).

[14] Gebremedhin, B. and Hoekstra, D., 2008. *Cereal marketing and household market participation in Ethiopia: the case of teff, Wheat and rice* (No. 307-2016-4963, p. 5).

[15] Gopalan, C., Ramashastry, B. V. and Balasubramanian, S. C., 1994. Nutritive Value of Indian Foods, Indian Council of Medical Research, pp. 24-26.

[16] Harer, P. N. and Deshmukh, R. B. 1992. Genetic variability, correlation and path coefficient analysis of soybean (Glycine max (L.) Merrill). Journal of Oilseed Research 9 (1): 65-71.

[17] Harland, S. C., 1939. The genetics of cotton jonathan cape, London. HAYMAN, BD and MATHER 1955. The description of genetic interaction in continuous variation. *Biometrics*, 11: 68-82.

[18] Hina Kausar, 2006. Genetic investigations in segregating populations of soybean (Glycine max (L.) Merrill). Karnataka Journal of Agricultural Sciences 19 (1): 200.

[19] JARC (Jimma Agricultural Research Center). 2014. Ethiopian Agricultural research, institute Jimma agricultural center profile report pp 5-7.

[20] Johnson, H. W., Robinson, H. F. and Comstock, R. E., 1955. Genotypic and phenotypic correlations in soybeans and their implications in selection 1. *Agronomy journal*, *47* (10): 477-483.

[21] Kumar. M., Karnwal, M. k. and Kamendra S., 2009. Studies on 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.

[22] Mahalanobis, P. C., 1936. On the generalized distance in statistics. National Institute of Science of India 2: 49-55.

[23] Mehetre, S. S., Mahajan, C. R., Desai, N. S. and Shinde, R. B., 1995. Variability, heritability and character association in M3 families of gamma irradiated soybean. *Soybean Genetics Newsletter*, 22: 126-131.

[24] MoFA and CSIR., 2005. Soybean Production Guide. Food crops development project. Ghana's Ministry of Food and Agriculture. 38pp.

[25] Mohana Rao, K. V., 1999. Genetic divergence in soybean (Glycine max (L.) Merrill). M. Sc. (Ag) Thesis, ANGRU Agricultural University, Hyderabad.

[26] Mpepereki, S., Javaheri, F., Davis, P. and Giller, K. E., 2000. Soybeans and sustainable agriculture: promiscuous soybeans in southern Africa. *Field crops research*, *65* (2-3): 137-149.

[27] Ramana, M. V., 2003. *Genetic Studies On Soybean (Glycine Max (L.) Merrill) In Non-Traditional Areas and Seasons* (Doctoral dissertation, Acharya NG Ranga Agricultural University; Hyderabad).

[28] Ramana, M. V., Rani, B. P. and Satyanarayana, A., 2000. Genetic variability, correlation and path analysis in soybean. *Journal of Oilseeds Research*, *17* (1): 32-35.

[29] Ramteke, R., Vineet, K., Pooja, M. and Agarwal, D. K., 2010. Study on genetic variability and traits interrelationship among released soybean varieties of India [Glycine max (L.) Merrill]. *Electronic Journal of Plant Breeding*, *1* (6): 1483-1487.

[30] SAS, S., 2012. STAT 9.3 User's guide. *Cary, NC: SAS Institute Inc*.

[31] Shurtleff, W. and Aoyagi, A., 2007. The soybean plant: Botany, nomenclature, taxonomy, domestication and dissemination. *Soy Info Center*.

[32] Siczek A. and Lipiec J. 2009. Soybean nodulation and symbiotic nitrogen fixation in response to soil compaction and mulching. EGU General Assembly 11 (1): 7520.

[33] Singh, R. K., 1985. Biometrical methods in Quantitative Genetic Analysis. Kalyani Pub. Ludhiana. *New Delhi, Revised Ed.*, *318*.

[34] Tesfaye, A., Githiri, M., Derera, J. and Debele, T., 2017. Genetic variability in soybean (Glycine max L.) for low soil phosphorus tolerance. *Ethiopian Journal of Agricultural Sciences*, *27* (2): 1-15.

[35] Wijnands J. H. M., Gurmesa N. D., Lute J. C. M. & Van Loo E. N. 2011. Ethiopian soybean and sunflower value chains: Opportunities and challenge. 2011-016.