American-Eurasian J. Agric. & Environ. Sci., 15 (10): 1927-1933, 2015
ISSN 1818-6769

DOI: 10.5829/idosi.aejaes.2015.15.10.12777

Genetic Divergence Analysis on Some Soybean (*Glycine max* L. Merrill) Genotypes Grown in Pawe, Ethiopia

[1]Tadesse Ghiday and [2]A. Sentayehu

[1]Pawe Agricultural Research Center, Ethiopia
[2]College of Agriculture and Veterinary Medicine, Jimma University, Ethiopia

Abstract: An investigation was carried out with 49 soybean genotypes to assess the diversity for yield and yield related traits. D-square statistics (D^2) has been used to classify the divergent genotypes into different groups. The genotypes were evaluated for 13 characters and showed moderate variability for the components studied. The cluster analysis grouped the 49 soybean genotypes into five different clusters. This indicates the presence of moderate diversity among the tested genotypes. From cluster mean values, genotypes in cluster III and V deserve consideration for their direct use as parents in hybridization programs to develop high yielding soybean varieties. The results of the principal component analysis revealed that five principal components (PC1 to PC5) accounted nearly for 79.06% of the total variation. The differentiation of the genotypes into different clusters was because of relatively high contribution within the first principal components such as number of pods per plant, biological yield, seed yield per plot and seed yield per plant. Therefore, the above mentioned characters which load high positive contribution more to the diversity and they were the ones that most differentiated the clusters. It was also noted that differentiation of genotypes into different clusters was because of the small contribution of few characters rather than the cumulative effect of a number of characters. The information obtained from this study can be used to plan crosses and maximize the use of genetic diversity and expression of heterosis.

Key words: Soybean · Hybridization · Genotypes · Genetic divergence · Ethiopia

INTRODUCTION

Soybean [*Glycine max* (L.) Merrill] is the most important vegetable food sources in the world. In Ethiopia, soybean is an introduced crop and had a higher expansion of cultivated area in recent years, with a crop production of 636531.01quintal of harvest with an average of productivity 19.98 quintal per hectare in 2012/2013 cropping season [1]. National average yield is very low compared with its potential and yields obtained in other soybean producing countries. It is largely grown in the lowlands of the country and constitutes roughly 2-3% of the annual pulse production and plays an appreciable role in human nutrition and health, edible oil, livestock feed and many other industrial and pharmaceutical applications [3].

For a successful breeding program, the presence of genetic diversity and variability play a vital role. Genetic diversity is essential to meet the diversified goals of plant breeding such as breeding for increasing yield, wider adaptation, desirable quality and pest disease resistance [6]. Genetic divergence analysis estimates the extent of diversity existed among selected genotypes [9]. Precise information on the nature and degree of genetic diversity helps the plant breeder in choosing the diverse parents for purposeful hybridization [10].

Genetic diversity plays an important role in plant breeding either to exploit heterosis or generate productive recombinants. The choice of parents is of paramount importance in breeding programs. Thus, the knowledge of genetic diversity and relatedness in the germplasm is a pre-requisite for crop improvement programs. Reduction in the genetic variability makes the crops increasingly vulnerable to diseases and adverse climatic changes [12]. So, precise information on the nature and degree of genetic diversity present in soybean introductions from principal areas of cultivation would help to select parents for evolving superior varieties. The aim of this study was to identify genetically divergent soybean parents with desirable traits for hybridization particularly for yield.

Corresponding Author: Tadesse Ghiday, Pawe Agricultural Research Center, P.O. Box 25, Pawe, Ethiopia.
Tel: +251(0)585500271.

MATERIALS AND METHODS

Forty-nine soybeans nationally released and introduced varieties were used in the experiment (Table 1). The experiment was conducted at Pawe Agricultural Research Center (PARC), North West Ethiopia from 2013 to 2014 summer season in a simple lattice design with two replications, each plot with four rows of 0.40m width and 5 m row length. Sowing was done by hand drilling at a seed rate of 70 kg/ha. The spacing between plots and replication were 0.40m and 1 m, respectively. Di ammonium phosphate (DAP) fertilizer was applied at the rate of 100 kg/ha. All the cultural practices were performed as recommended. The plant data during the cropping season and after harvesting were recorded. Observations recorded on a plot basis included days to flowering, days to maturity, seed filling period, 100 seeds weight, biological yield per plot, seed yield per plot and harvest index per plot. The seed yield per plot and biological yield per plot were measured by taking a net plot of 0.8 m x 5 m or 4 m^2 and this was used to determine harvest index. All other characters were recorded on a single plant basis by randomly taking five plants from each experimental plot. Climatic conditions for the cropping seasons have been indicated in Fig. 1 and 2.

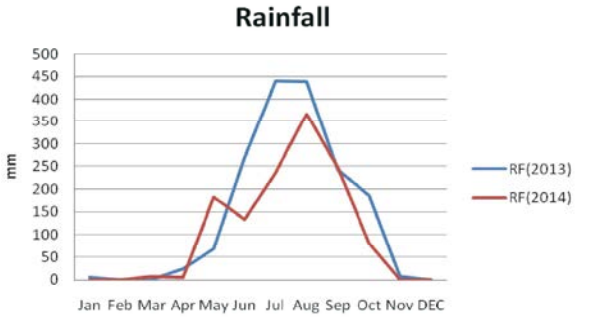

Fig. 1: Monthly total rain fall (mm) of Pawe Research Center, 2013 and 2014

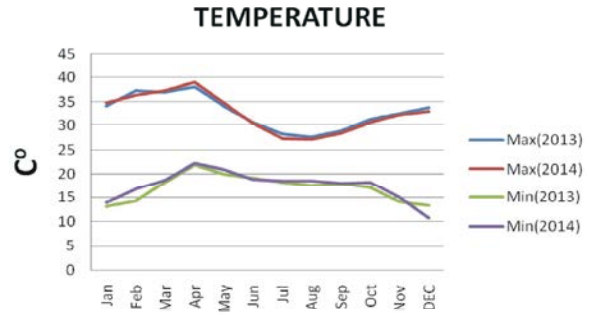

Fig. 2: Monthly average maximum and minimum temperatures (C°) of Pawe Research Center, 2013 and 2014.

Table 1: Forty-nine soybean genotypes of different crosses (hybrids), nationally released and introduced varieties used in this study (2013/2014)

Entry No	Genotype	Source	Year	Entry No	Genotype	Source	Year
1	AFGAT (TGX-1892-10F)	IITA/Nigeria	2007	26	TGX-1987-34F	IITA/Nigeria	2007
2	AWASSA 95 (G 2261)	USA	2005	27	TGX-1987-11F	IITA/Nigeria	2007
3	BLACK-HAWK	USA	2005	28	TGX-1740-2F	IITA/Nigeria	2007
4	CLARK 63K	USA	2005	29	TGX-1987-9F	IITA/Nigeria	2007
5	PROTONA-2	USA	2005	30	TGX-1987-23F	IITA/Nigeria	2007
6	NYALA	Turkey	2003	31	TGX-1987-64F	IITA/Nigeria	2011
7	ETHIO-YOGOSLAVIA	USA	2005	32	TGX-1987-62F	IITA/Nigeria	2011
8	WOGAYEN	USA	2005	33	TGX-1987-15F	IITA/Nigeria	2011
9	GIZO	USA	2005	34	TGX-1986-3F	IITA/Nigeria	2011
10	GISHAMA	USA	2005	35	TGX-1987-35F	USA	2011
11	AGS 7-1	USA	2005	36	TGX-1987-19F	IITA/Nigeria	2011
12	NOVA	USA	2005	37	TGX-1935-10E	IITA/Nigeria	2011
13	WELLO	USA	2005	38	TGX-1987-40F	IITA/Nigeria	2011
14	GOZILLA	USA	2005	39	TGX-1987-38F	IITA/Nigeria	2011
15	EAZ-3600	USA	2005	40	TGX-1987-37F	IITA/Nigeria	2011
16	BELESSA 95 (PR-149)	USA	2005	41	TGX-1987-14F	IITA/Nigeria	2011
17	LOTTUS	USA	2005	42	TGX-1987-10F	IITA/Nigeria	2011
18	PARC-1	USA	2005	43	TGX-1987-65F	IITA/Nigeria	2011
19	PARC-2	USA	2005	44	CROWFORD	IITA/Nigeria	2007
20	PARC-3	USA	2005	45	WILLIAMS	IITA/Nigeria	2007
21	PARC-4	USA	2005	46	COCKER-240	IITA/Nigeria	2007
22	PARC-5	USA	2005	47	BOSHE	IITA/Nigeria	2007
23	PARC-6	USA	2005	48	JALELE	IITA/Nigeria	2007
24	TGX-1987-18F	IITA/Nigeria	2007	49	TGX-1989-59F	IITA/Nigeria	2007
25	TGX-1987-20F	IITA/Nigeria	2007				

Source: Pawe Agricultural Research Center 2011

Statistical Analysis: The statistical package SAS version 9.2 was used for genetic divergence calculation, Cluster mean analysis and principal component analysis (SAS Institute, 2008).

RESULTS AND DISCUSSION

Analysis of Variance: Mean squares of the 13 characters from analysis of variance (ANOVA) are presented in Table 2. Significant differences among genotypes ($P<0.001$) were observed among the 49 genotypes for 13 of the traits studied, indicating the presence of inherent variation among the materials. Desirable genes from this germplasm can effectively be utilized to develop high performing pure line varieties after crossing. The present study are in agreement with those obtained by Bangar [4] on soybean.

Genetic Divergence: Differences in morphological and quantitative traits have been considered as simple indicator of genetic variability in crop species and varieties. Divergence analysis is a technique used to categorize genotypes that are similar as possible into one group and the other into different. D-square statistics (D^2) developed by Mahalanobis [8], has been used to classify the divergent genotypes into different groups. The extent of diversity present between genotypes determines the extent of improvement gained through selection and hybridization. The more divergent the two genotypes are the more will be the probability of improving through selection and hybridization.

The D-square statistics (D^2) resulted in classifying the 49 soybean genotypes into five distinct clusters (Table 3 and Fig. 3). This indicates the presence of moderate diversity among the tested genotypes.

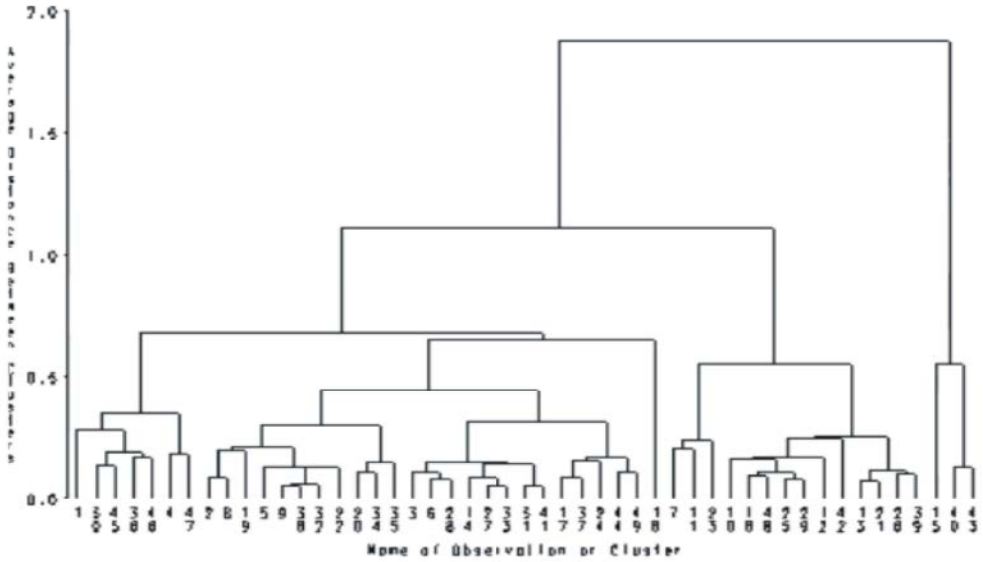

Fig. 3: Figure showing the clusters to which the genotypes belong and average distance between clusters (2013 and 2014).

Table 3: The distribution of genotypes into 5 clusters based on D^2 analysis for 49 soybean genotypes tested at Pawe (2013 and 2014)

Cluster	No of genotypes	Percentage (%)	Genotypes
I	24	48.98	GIZO, GISHAMA, TGX1987-11F, TGX1987-15F, TGX1987-64F, TGX1987-14F, TGX1987-62F, NYALA, TGX1740-2F, LOTUS, TGX1935-10E, AWASSA-95, WEGAYEN, GOZELLA, CROWFORD, TGX1989-59F, BLACK HAWK PARC-3, TGX1986-3F, PARC-5, PROTONA-2, TGX1987-35F, TGX1987-18F and PARC-2
II	14	28.57	WELLO, PARC-4, TGX1987-20F, TGX1987-9F, PARC-1, JAKELE, TGX1987-34F, TGX1987-38F, GISHAMA, NOVA, TGX 13-3-2644, AGS 7-1, PARC-6 and TGX1987-10F
III	3	6.12	TGX1987-37F, TGX1987-10F and EAZ-3600
IV	7	14.29	TGX1987-23F, WILLIAMS, TGX1987-19F, COKER-240, CLARK-63K, BOSHE and AFGAT
V	1	2.01	BELESA-95,

Table 4: Mean value of 13 characters for the 5 clusters of 49 soybean genotypes tested at Pawe (2013 and 2014)

Character	Cluster I	Cluster II	Cluster III	Cluster IV	Cluster V
DF(days)	61.08	61.04	64.17**	54.79	49.5*
DM(days)	113.77	116.90**	108.67	115.93	96*
PH(cm)	58.60	71.76**	66.43	55.11	44.6*
SFP(days)	52.69	55.86	44.5*	61.14**	46.5
NBP(N°)	2.35	2.64	2.78	2.23*	3.45**
NPP(N°)	15.49	20.50	22.18	12.35*	33**
NSP(N°)	2.41	2.52	2.67**	2.34*	2.55
BY(g)	859.94	1356.4	1972.87**	527.99*	1434.57
SYP(g)	264.99	367.77	386.98	166.6*	566.8**
SCAH(N°)	91.06	97**	83.5	81.57	75.5*
HSW(g)	12.62	11.79	11.43*	12.97**	12.6
HI(%)	30.52	27.33	19.44*	31.07	39.51**
SY(g)	2.89	3.80	4.77	2.15*	7.51**

* and **low and high value of a trait respectively

DF= days to flowering, DM=days to maturity, SFP= seed filling period, PH=plant height, NBP=number of branches per plant, NPP= number of pods per plant, NSP= number of seeds per plant, BY= biological yield, SYP= seed yield plot, SCAH= stand count at harvest, HSW= hundred seed weight, HI= harvest index and SY=seed yield per plant

Table 4: Pair wise generalized squared distance (D^2) among 49 soybean genotypes in five clusters at Pawe (2013 and 2014)

	Generalized mean square distance				
Cluster	I	II	III	IV	V
I		23.67169*	163.7231**	16.72047^{ns}	181.2203**
II			82.2458**	75.68003**	220.8855**
III				250.7112**	252.604**
IV					172.9535**
V					

x^2= 21.03 and 26.22 at 5%, 1% probability level respectively

•**and**significant and highly significant at5% and1% probability level respectively*

Table 5: Eigen vectors and Eigen values of the first five principal components of 49 soybeans genotypes evaluated at Pawe (2013 and 2014)

	Principal component analysis				
Trait	PC 1	PC 2	PC 3	PC 4	PC 5
DF	-0.042	0.341	-0.223	0.139	-0.239
DM	-0.011	0.423	0.320	0.349	0.080
SFP	0.014	0.264	0.504	0.308	0.242
PH	0.178	0.439	0.010	0.153	0.014
NBP	0.226	-0.008	-0.098	0.339	-0.352
NPP	0.473	-0.135	0.066	0.073	0.024
NSP	0.076	-0.139	-0.326	0.156	0.514
BY	0.350	0.101	-0.356	0.149	0.272
SYP	0.501	-0.020	0.0335	-0.092	0.007
SCAH	0.145	0.403	-0.069	-0.474	0.032
HSW	0.002	-0.130	0.351	-0.217	0.543
HI	0.227	-0.166	0.453	-0.241	-0.341
SY	0.471	-0.182	0.060	0.078	-0.039
Eigen Value	3.340632	2.903523	1.992596	1.648593	1.183332
Difference	0.437109	0.910926	0.344003	0.465261	0.216324
Proportion	0.2386	0.2074	0.1423	0.1178	0.0845
Cumulative	0.2386	0.446	0.5883	0.7061	0.7906

DF= days to flowering, DF=days to maturity, GFP= grain filling period, PH=plant height, NBP=number of branches per plant, NPP= number of pods per plant, NSP= number of seeds per plant, BY= biological yield, SYP= seed yield plot, SCAH= stand count at harvest, HSW= hundred seed weight, HI= harvest index and SY=seed yield per plant

The cluster analysis based on the pooled mean of genotypes resulted in classifying the 49 genotypes into four groups and one solitary (Table 3). This indicates that tested soybean genotypes were moderately divergent. The chi-test for the five clusters indicated that there was statistically accepted difference between clusters (Table 4). The genotypes were distributed (Table 3) in such a way that 24 genotypes were grouped into Cluster-I (49.98%), 14 genotypes in to Cluster-II (28.57%), 3 genotypes into Cluster-III (6.12%), 7 genotypes into Cluster-IV (14.29%) and one genotype into Cluster-V (2.01%). Cluster-I contains moderate value of characters. Cluster-II contains high mean number of days to maturity (116.90days), stand count at harvest (97) and plant height (71.76cm). Cluster-III contains high days to flowering (64.17days), number of seeds per pod (2.62) and biological yield (1972.87). In Cluster-III there were characters such as seed filling period (44.5days), hundred seed weight (11.43g) and harvest index (19.44 %) of lower value. Cluster-IV contains high seed filling period (61.14days) and hundred seeds weight (12.97g). Cluster IV contains low number of branches per plant (2.23), number of seeds per pod (2.34), biological yield (527.99g), seed yield per plot (166.6g), number of pods per plant (12.35) and seed yield per plant (2.15g). Cluster V contains high number of branches per plant (3.45), number of pods per plant (33.00), harvest index (39.51) and seed yield per plant (7.51). In Cluster V there were characters such as days to flower (49.5days), days to maturity (96days), plant height (44.6cm) and stand count at harvest (75.5) of lower value. The maximum inter cluster was between Cluster-III and V (252.6) followed by Cluster III and IV (250.71) and Cluster II and V (220.89) (Table 2). The minimum being Cluster I and IV (16.72) followed by Cluster I and II (23.67). Generally, this study showed that the genotypes included in this study are moderately divergent. Therefore, the results of the distance of genotypes between clusters has shown that there is a room for the genetic improvement of soybean varieties and the information generated can be used to plan wide crosses, to exploit genetic diversity and maximize the expression of hetrosis [6].

Cluster III and V, Cluster III and IV and Cluster II and V exhibited the greatest inter cluster divergence from all other cluster in this study. According to Gemechu and Ghaderi [5,7], increasing parental distance implies a great number of contrasting alleles at the desired loci and to the extent that these loci recombine in the F_2 and F_3 generation following a cross of distantly related parents, the greater will be the opportunities for the effective selection for yield factors. Thus, crossing of genotypes from these clusters with other clusters may produce higher amount of hetrotic expression in the first filial generations (F_1's) and wide range of variability in subsequent segregating (F_2) populations. Thus, crosses involving cluster III and V with any other cluster is suggested to exhibit high heterosis and could result in segregates with higher seed yield, i.e. transgressive segregation.

Principal Component Analysis: Principal component analysis (PCA) reflects the importance of the largest contributor to the total variation at each axis of differentiation [11]. The Eigen values are often used to determine how many factors to retain. The sum of the Eigen values is usually equal to the number of variables. Therefore, in this analysis the first factor retains the information contained in 3.341 of the original variables. The principal components of these data are shown in Table 5.

Five principal components PC1 to PC5 which are extracted from the original data and having latent roots greater than one accounted nearly 79.06% of the total variation (Table 5), suggesting that these principal component scores might be used to summarize the original 13 variables in any further analysis of the data. Out of the total principal components retained, PC1, PC2, PC3, PC4 and PC5 with values of 23.86%, 20.74%, 14.23%, 11.78% and 8.45%, respectively contributed more to the total variation. According to Chahal [2] characters with largest absolute value closer to unity within the first principal component influence the clustering more than those with lower absolute value closer to zero. Therefore, in the present study, differentiation of the genotypes into different clusters was because of relatively high contribution of few characters rather than small contribution from each character.

In Table 5, the first principal component had high positive component loading from seed yield per plot (0.501), number of pods per plant (0.473), seed yield (0.471) and biological yield (0.350). The positive loading shows the presence of positive correlation trends between the components and the variables. Therefore, the above mentioned characters which load high positive contributed more to the diversity and they were the ones that most differentiated the clusters. The major contributing characters for the diversity in the second principal component (PC2) have high positive component loading from plant height (0.439), stand count at harvest

(0.403) and days to flowering (0.341). The major contributing characters for the diversity in the third principal component (PC3) had high positive component loading from seed filling period (0.504), harvest index (0.453), hundred seed weight (0.351) and days to maturity (0.320); and negative loading from biological yield (-0.356) and number of seeds per pod (-0.326). In principal component four (PC4) high positive component loading from days to maturity (0.349), number of branches per plant (0.339) and seed filling period (0.308) and high negative loading from stand count at harvest (-0.474). In principal component five (PC5) high positive component loading from hundred seed weight (0.543) and number of seeds per pod (0.514) and high negative loading from number of branches per plant (-0.352) and harvest index (-0.341). The positive and negative loading shows the presence of positive and negative correlation trends between the components and the variables. Therefore, the above mentioned characters which load high positively or negatively contributed more to the diversity and they were the ones that most differentiated the clusters.

Usually it is customary to choose one variable from these identified groups. Hence, for the first group grain yield per plot (0.501) is best choice, which had the largest loading from component ones, plant height (0.439) for the second, seed filling period (0.504) for the third group, stand count at harvest (-0.474) for the fourth group and hundred seed weight (0.543) for the fifth group.

SUMMARY AND CONCLUSION

There were differences in the performance of the genotypes as there were statistically supported significant differences among genotypes for most of the 13 characters and relatively wide range of the mean values for most of the characters. Nevertheless, the level of the genetic differences for many traits, including grain yield, may not be sufficient to expect progress in selection. Therefore, in order to improve the diversity of soybean in Ethiopia, subsequent crossing program aimed at developing soybean varieties of better diversity by crossing between highly divergent varieties needs to be carried out.

The cluster analysis based on D^2 analysis on pooled mean of genotypes classified the 49 genotypes into five clusters, which makes them to be moderately divergent. There were statistically significant differences between all of the clusters, except between cluster I and IV.

The principal component analysis extracted five principal components PC1 to PC5 from the original data and having Eigen value greater than one accounting nearly 79.06% of the total variation. Characters with largest absolute value closer to unity within the first principal component such as number of pods per plant, biological yield, seed yield per plot and seed yield per plant influence the clustering. The differentiation of the genotypes into different clusters was because of relatively high contribution of these characters. Therefore, the above mentioned characters which load high positive contributed more to the diversity and they were the ones that most differentiated the clusters. The present investigation provided considerable information useful in genetic improvement of soybean. Genotype grouped into cluster V showed maximum inter cluster diversity. From cluster mean values, genotypes in cluster III and V deserve consideration for their direct use as parents in hybridization programs to develop high yielding soybean varieties. There is significant genetic variability among tested genotypes that indicates the presence of better opportunity to bring about improvement through wide hybridization by crossing genotypes in different clusters. Further studies on the soybean genotypes have to be tested for more locations and seasons to recommend highly performed ones. More genotypes and more number of characters have to be included for sound recommendation. Crossing of genotypes from distant clusters enables to have more variability for desirable traits for improvement. Influence of environment and an agronomic practice on the genetic potential of the varieties in different soybean environments is necessary. This is helpful to stratify the environments based on quality and yield suitability. Generally, the development of soybean varieties possessing higher seed yield, higher protein and oil content is important.

Recommendation: Further studies on the soybean genotypes have to be tested for more locations and seasons to recommend highly performed ones. More genotypes and more number of characters have to be included for sound recommendation. Crossing of genotypes from distant clusters enables to have more variability for desirable traits for improvement. Influence of environment and an agronomic practice on the genetic potential of the varieties in different soybean environment is necessary. This is helpful to stratify the environments based on quality and yield suitability. Generally, the development of soybean varieties possessing higher grain yield, higher protein and oil content is important.

REFERENCES

1. FAO., 2012. FAOSTAT, FAO Statistical Data Bases-agriculture (available at http:// apps. fao. org.). http://faostat.fao.org/ FAOSTAT.
2. Chahal, G.S. and S.S. Gosal, 2002. Principles and Procedures of Plant Breeding: Biotechnology and Conventional Approaches. Narosa Publishing House, New Delhi.
3. CSA., 2012. Agricultural Sample Survey Statistical Bulletins. Central Statistical Authority, Addis Ababa, Ethiopia.
4. Bangar, N.D., G.R. Mukhekar, D.B. Lad and D.G. Mukhekar, 2000. Genetic variability, correlation and regression studies in soybean. J. Mah. Agri. Univ., 28(3): 320-321.
5. Gemechu, K., J. Musa, W. Tezera and D. Getnet, 2005. Extent and pattern of genetic diversity for morph-agronomic traits in Ethiopian highland pulse landraces: I. Feild Pea (*Pisum sativum* L.). Genetic Resources and Crop Evolution, 52: 539-549.
6. Gemechu, K., B. Endashaw, A. Mohammed, D. Kifle, G. Emana and A. Fassil, 2012. Genetic diversity and population structure of Ethiopian Chick Pea (*Cicer arietinum* L.) Germplasm accessions from different geographic origins as revealed by microsatellite markers, Plant Mol. Biol. Rep., 30: 654-665.
7. Ghaderi, A., M.W. Adams and A.M. Nassib, 1984. Relationship between genetic distance and heterosis for yield and morphological traits in dry edible bean and faba bean. Crop Sci., 24: 37-42.
8. Mahalanobis, P.C., 1936. On the generalized distance in statistics. Proc. Natl. Sci. India B., 2: 49-55.
9. Ramigiry, S.R., 1999. Genetic divergence in soybean. Madras Agricultural Journal, 8(3): 5167-170.
10. Sharma, S.S., 2005. Genetic divergence in Indian varieties of soybean. Soybean Research, 3: 9-16.
11. Sharma, J.R., 1998. Statistical and Biometrical Techniques in Plant Breeding. New Age International (P) Limited Publishers, New Delhi.
12. Punia, S.S., R. Baldev, N.R. Koli, B.R. Ranwah, P. Rokadia and S.R. Maloo, 2011. Genetic architecture of quantitative traits in field pea. Journal of Food Legumes, 24: 299-303.