Journal of Natural Sciences Research www.iiste.org
ISSN 2224-3186 (Paper) ISSN 2225-0921 (Online)
Vol.5, No.9, 2015

Grain Yield Based Cluster Analysis and Correlation of Agronomic Traits of Sesame (Sesamum indicum L.) Genotypes in Ethiopia

Fiseha Baraki[1*] Yemane Tsehaye[2] Fetien Abay[2]
1.Crop Research Core Process, Humera Agricultural Research Center, Ethiopia
2.Department of Crop and Horticultural Science, Mekelle University, Ethiopia
*Corresponding Author Email: fish051bar@gmail.com

Abstract

The experiment was conducted in the western Tigray region of Northern Ethiopia in three locations and three growing seasons (a total of seven environments). The objective of the study was to assess the interrelation ship of the genotypes in their grain yield and other agronomic traits. The experiment was laid out in RCBD with a total and harvestable plot size of 14m^2 and 10 m^2 respectively. In the experiment a total of thirteen sesame genotypes were used and evaluated for the interrelation ship of their grain yield and other agronomic traits. Based on their grain yield the thirteen genotypes were grouped in to four clusters and the maximum genetic divergence (D^2= 215.57) was observed between cluster-II and cluster-III. Grain yield had a significant and positive correlation ($P<0.01$) with oil content ($r = 0.79$), number of branches ($r = 0.7$) and number of capsules ($r = 0.68$). In contrast to this, yield was negatively correlated with days to maturity ($r = -0.58$) and days to flowering ($r = -0.11$). Exhaustive assessment of the association of grain yield and other agronomic traits is indispensably important for further breeding program and enhancing grain yield of sesame.

Keywords: Cluster, Correlation, Genotype, Grain yield

1. Introduction

Sesame (*Sesamum indicum* L.) which is also commonly known as til (Hindi), hu ma (Chinese), sesame (French), goma (Japanese), gergelim (Portuguese) and ajonjoli (Spanish) (Anilakumar, *et al*., 2010) is an annual plant that belongs to the Pedaliaceae family. Sesame is mainly cultivated between the 25° N and 25° S latitudes. Despite its ideal adaptation to dry sites, sesame can also be cultivated in humid, tropical and sub-tropical regions (Natureland, 2002). It is an erect herbaceous annual plant with either single stemmed or branched growth habits and two growth characteristics of indeterminate and determinate, reaching up to 2m height and with a large tap root of up to 90 cm (Ashri, 2007). According to Kobayashi, *et al.* (1990), 36 species have been identified under the genus *Sesamum* with three cytogenetic groups $2n = 26$, $2n = 32$ and $2n = 64$ from which the widely cultivated *Sesamum indicum* is within the first group. Ethiopia is the Sixth largest sesame producer in the world following Myanmar, India, China, Tanzania and Uganda, respectively and third in Africa preceding Nigeria and Mozambique, respectively (FAOSTAT, 2012). Within Ethiopia, the Tigray region is the main producer with large commercial farms and many small scale farmers and it is a good source of income in these areas using as a cash crop and local oil extraction.

Cluster analysis aims to classify a sample of subjects (or objects) on the basis of a set of measured variables into a number of different groups such that similar subjects are placed in the same group (Cornish, 2007). Ward's method is distinct from all other methods because it uses an analysis of variance approach to evaluate the distances between clusters and it attempts to minimize the Sum of Squares (SS) of any two clusters that can be formed at each step. Observations in a specific cluster share many characteristics in common, but are very dissimilar to observations not belonging to that cluster (Ward, 1963). Among the different types of clustering hierarchical clustering is by far the most widely used clustering method. In cluster analysis performing Mahalanobis's (D^2) statistic analysis Mahalanobis (1936) is important which uses for assessing the genetic divergence among the test entries in different clusters. If there is a large jump in the distance between clusters from one stage to another then this suggests that at one stage clusters that are relatively close together were joined whereas, at the following stage, the clusters that were joined were relatively far apart.

In any plant breeding program the final objective is to boost quality and/or quantity of a required crop. knowing the association between the required traits and other related traits is a prerequisite for such programs. Hence, correlation coefficient can measure such simple linear relationship between different traits. Correlation coefficients range between -1 and +1 and measure the degree and direction of relationships between various traits. Yield is a dependable complex inherited character as a result of interaction of several contributing factors that may be related or unrelated (Subramanian and Subramanian, 1994).

2. Material and Methods

2.1 Description of the Study Area

The experiment was conducted in Northern Ethiopia, Tigray region (specifically Humera, Dansha and Sheraro). Edaphic and climatic description of the specific study areas is listed in table 1 in addition to the geographic location depicted in figure 1.

Table 1: Agro-climatic and soil characteristics of the experimental sites

Location	Latitude (^{o}N)	Longitude (^{o}E)	Altitude (m)	Anual RF (mm)	Min - Max Temp (^{o}c)	Soil texture		
						Clay (%)	Silt (%)	Sand (%)
Humera	14^{o}15'	36^{o}37'	609	576.4	18.8-37.6	35.6	25.6	38.6
Sheraro	14^{o}24'	37^{o}45'	1028	676.7	18.8-34.9	21	27.3	51.7
Dansha	13^{o}36'	36^{o}41'	696	888.4	28.7(mean)	_	_	_

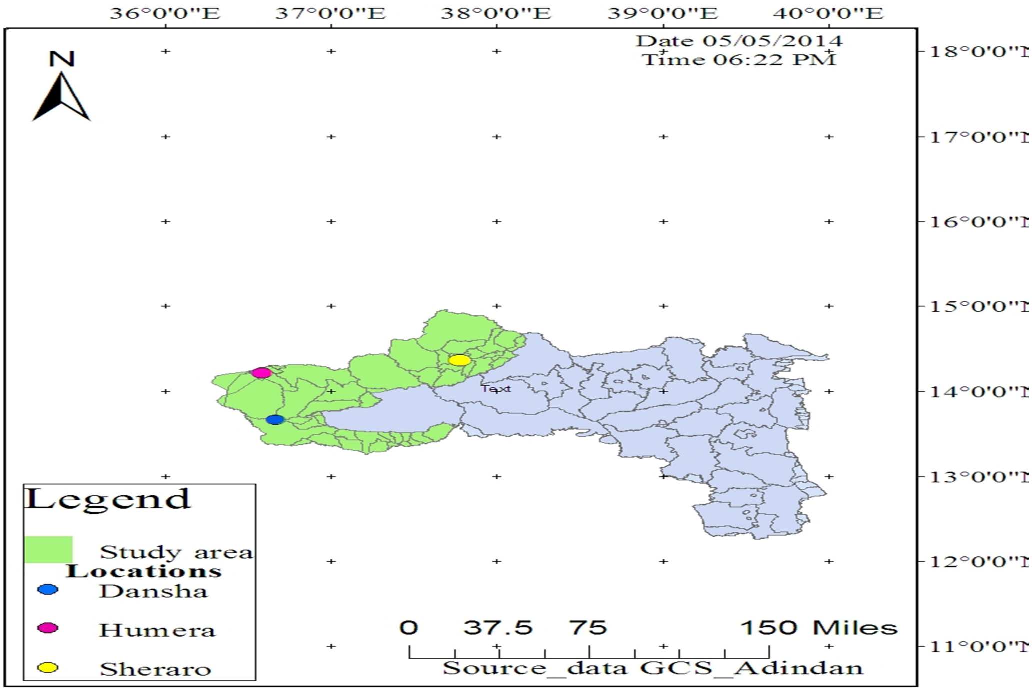

Figure 1: Geographic location of Tigray and the specific study areas

2.2 Experimental Material and Method

The experiments were conducted for three growing seasons (2011-2013) in Humera and Dansha, and in a single year (2013 cropping season) in Sheraro (total seven environments), in Tigray region, Ethiopia, under rain fed condition. where: E1, E2, E3 are 2011, 2012, 2013 growing seasons respectively in Humera; E4, E5, E6 are 2011, 2012, 2013 growing seasons respectively in Dansha; and E7 is 2013 growing season in Sheraro. In detail description of the genotypes is listed in table 2. Thirteen sesame genotypes were sown in RCBD with three replications and evaluated for the interrelation ship of their grain yield and other agronomic traits. Each plot had a total of seven rows with five harvestable ones. Each genotype was randomly assigned and sown in a plot area of 2.8 m by 5m with 1m between plots and 1.5 m between blocks keeping inter and intra row spacing of 40 cm and 10 cm, respectively.

Table 2: Description of the sesame genotypes

Genotype name	Gen code	Status	Seed color	Source
Acc#031	G1	Advanced line	White	WARC
Oro (9-1)	G2	Advanced line	White	WARC
NN-0079-1	G3	Advanced line	White	WARC
Acc-034	G4	Advanced line	White	WARC
Abi-Doctor	G5	Advanced line	White	WARC
Serkamo	G6	Released	Brown	WARC
Acc-051-020sel-14	G7	Advanced line	Brown	WARC
Tate	G8	Released	Brown	WARC
Acc-051-02sel-13	G9	Advanced line	White	WARC
Adi	G10	Released	White	WARC
Hirhir	G11	Farmers seed (local check)	White	HuARC
Setit-1	G12	Released (standard check)	White	HuARC
Humera-1	G13	Released (standard check)	White	HuARC

WARC-Werer Agricultural Research Center, HuARC-Humera Agricultural Research Center

2.3 Data Collection

From a net plot area ten plants were selected randomly and tagged to collect the agro-morphological data such as, plant height, length of capsule bearing zone, number of branches and number of capsules. The average values of ten plants were considered for further analysis. The seeds from each plot were weighed for yield determination. Oil content was determined from the composite of the three plots of each genotype. The detail description of the phenological, yield and yield component traits is described below.

- ✓ ***Days to 50% flowering (DF)***: The number of days from emergence to which 50% of the population in each plot become flowered.
- ✓ ***Days to 75% maturity (DM)***: The number of days from emergence to when 75% of the plants in each plot had fully matured.
- ✓ ***Plant height at maturity (PH) (cm)***: This growth parameter was measured from ten randomly selected and tagged plants from the harvestable rows of each plot with the help of meter tape from ground surface to the top of the plant.
- ✓ ***Length of capsule bearing zone (LCBZ) (cm)*** : A height from the first capsule to tip of the plant, measured using meter tape.
- ✓ ***Number of primary branches per plant (NB)***: Branches producing productive capsules will be recorded for randomly selected plants.
- ✓ ***Number of capsules (NC)***: The total number of capsules was counted from ten randomly selected plants at maturity.
- ✓ ***Grain yield (kg/ha)***: the total grain yield harvested from the net plot area was weighed using a sensitive balance.
- ✓ ***Oil contents (OC) (%)***: Oil content was determined by wide line nuclear magnetic resonance (NMR) and it was explained in percentage.

2.4 Statistical Analysis

Statistical estimations and computations were performed using different statistical software. Homogeneity of residual variances was tested prior to a combined analysis over locations in each year as well as over locations and years using Bartlet's test (Steel and Torrie, 1998). Accordingly, the data collected were homogenous and all data showed normal distribution.

2.4.1 Association of Genotypes and their Characters

Using the mixed model and considering the genotype effect as random and replication as fixed, the Best Linear Unbiased Predictor (BLUP) was estimated via REML (restricted maximum likelihood) for each genotype in each environment. The major advantage of BLUP is shrinking towards the mean (Piepho, *et al.*, 2008) and it maximizes the correlation of true genotypic values and predicted genotypic values (Searle, *et al.*, 1992). The correlation, clustering and Mahalanobis (D^2) distance was executed using the BLUP values of the genotypes.

Cluster analysis based on Ward's method (Ward 1963) using squared Euclidean distance of the distance metric and standardized variables was performed using Minitab release 16 (Minitab, 1998) to cluster the genotypes based on their grain yield in different environments Moreover, Genetic divergence among the thirteen genotypes was estimated using the Mahalanobis (D^2) statistics (Mahalanobis, 1936). The D^2 values obtained from pairs of clusters were considered as the calculated values of Chi-square ($\chi 2$) and were tested for significance at 1% and 5% probability level against the tabulated values of $\chi 2$ for 'P' degree of freedom, where P is the number of characters considered (P=8) (Urdan, 2005). Furthermore, to determine the strength of relationships of grain yield with the other traits and even among each other of the other traits Pearson's correlation coefficient was determined using the Minitab release 16 (Minitab, 1998).

3. Result and Discussion

3.1 Grain Yield clustering of Sesame Genotypes

Cluster analysis is a tool that creates subgroups that are more manageable than individual datum and it examines the full complement of inter-relationships between variables. The genotypes were clustered in four groups based on their grain yield in the different environments (Figure 2). In addition the mean grain yield of the clusters in each environments and the grand mean of grain yield for each cluster is illustrated in table 3 . Cluster I comprised five genotypes viz. G5, G2, G7, G3 and G6 with an average yield of 691.2 kg/ha (Table 3). The exceptionally promising genotypes, G1 and G4, with higher average grain yield (899.42 kg/ha) were in cluster II. Cluster III, the lowest yielding cluster (661.84 Kg/ha), encompassed three genotypes, G8, G9 and G10 within it. The varieties G11, G 12 and G13, with an average yield of 805 kg/ha were grouped in cluster IV. Parsaeian, *et al.* (2011) used cluster analysis and grouped different genotypes of sesame in to five similar clusters based on their Agro-morphological traits. Likewise, Tabatabaei, *et al.* (2011) and Tripathi, *et al.* (2013) also clustered different sesame genotypes based on their morpho-physiological traits, and morphological and quality traits correspondingly.

To validate the observed grouping using the cluster analysis, distance among clusters was assessed using the Mahalanobis's (D^2) distance (Table 4). The distance among clusters were statistically significant. The minimum and maximum distance was between clusters II and IV ($D^2 = 20.155$), and II and III ($D^2=215.573$) respectively. The distance between clusters II (the promising genotypes) and III (the poorly yielding genotypes) was extremely larger than any clusters ($D^2= 215.573$) (Table 4). Mahalanobis D² statistic is more reliable in selection of potential parent for hybridization program using these D² values cluster. From this point of view the members of cluster II , G1 and G4, which are with higher grain yield may be used as parent materials for further breeding program to improve grain yield of sesame. Singh *et al.* (2005) and More *et al.* (2006) were also used Mahalanobis D² to determine the genetic divergence of different maize genotypes.

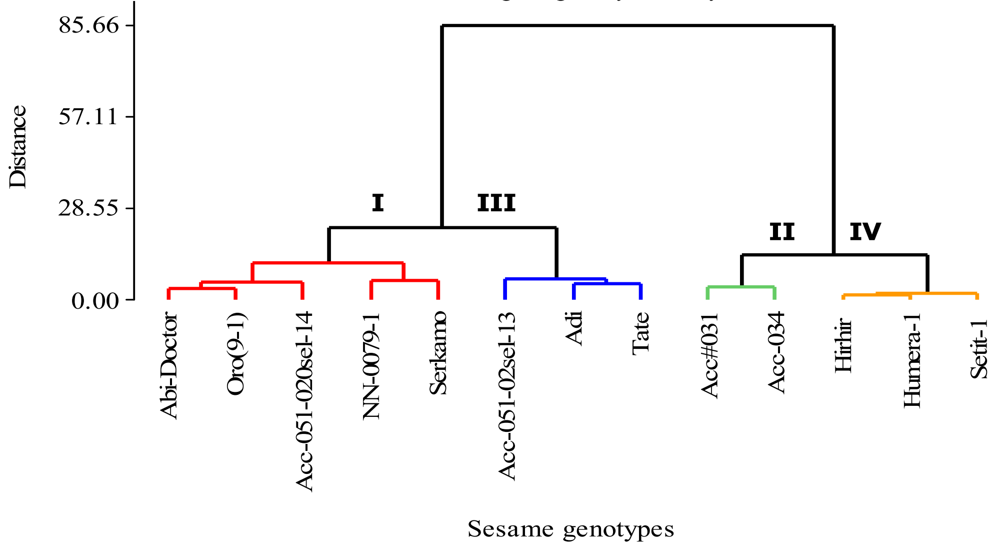

Figure 2: Dendrogram of sesame genotypes showing the degree of their relationship in grain yield

Where: Acc# 031 (G1), Oro (9-1) (G2), NN-0079-1 (G3), Acc-034 (G4), Abi-doctor (G5), Serkamo (G6), Acc-051-02sel-13 (G7), Tate (G8), Acc-051-020sel-14 (G9), Adi (G10), Hirhir (G11), Setit-1 (G12), and Humera-1 (G13)

Table 3. Mean grain yield of sesame genotypes over the different environments

	Environments (E)							
Cluster	**E1**	**E2**	**E3**	**E4**	**E5**	**E6**	**E7**	**Cluster Mean**
I	614.7	570.3	741.4	601.6	738.1	811.6	760.5	691.2
II	833.9	884.0	824.6	856.2	875.8	978.5	1043.1	899.4
III	491.8	536.5	592.2	642.0	703.7	930.3	736.4	661.8
IV	650.5	786.8	817.0	795.1	822.7	935.6	832.4	805.7

Where: E1(Environment-1), E2(Environment-2), E3(Environment-3), E4(Environment-4), E5(Environment-5), E6(Environment-6), E7(Environment-7)

Table 4: Mahalanobis's (D^2) distance among grain yield based clusters

	I	II	III	IV
I	0			
II	101.891**	0		
III	104.535**	215.573**	0	
IV	46.475**	20.155**	173.205**	0

**Significant (P<0.01)

3.2 Correlation of Grain yield and Other Agronomic Traits

Correlation coefficient analysis was done for yield and other traits based on the average BLUP values estimated (considering genotypes as random and replication as fixed) for each environment. Grain yield had a significant and positive correlation (P<0.01) with number of branches (r = 0.7) and number of capsules (r =0.68). Results of Ahmed and Ahmed (2012), and Abd El-Mohsen (2013) revealed that sesame grain yield had a significant and positive relationship with number of branches and number of capsules. Similarly, (Aremu, *et al*., 2011; Pham 2011) found a positive and high correlation among number of capsules and grain yield of sesame. Furthermore, grain yield was significantly and positively correlated (r = 0.798) with oil content which is similar to the results of Chowdhury, *et al*. (2010) who, reported a positive correlation between grain yield and oil content. Grain yield had a non significant and weak correlation with PH (r = 0.22) and LCBZ (r = 0.33). The traits which showed positive association with grain yield had their own contribution in having better yield of the genotypes. As grain yield is a complex trait and is highly influenced by genetic as well as environmental factors, direct selection for yield may not be as such effective. Thus, selection based on capsule number and number of branches (as observed from the high bi-variate correlation in table 5) could enhance grain yield in sesame. For example, genotypes which scored the highest yield such as G1 and G4 had the highest branch as well as capsule number (data not shown), which indicates the positive and direct contribution of the traits for yield increment. On the contrary, yield was negatively correlated with days to maturity (r = -0.58) and flowering time (r = -0.11), but the association was statistically significant (P<0.05) with days to maturity and statistically significant (P<0.05) with days to flowering (Table 5). This is in agreement with the finding of Ong'injo (2009) and Yol (2010) who found weak and negative correlation of grain yield with days to 50% flowering but positively and highly correlated with number of branches and capsule. The negative association between yield and the phenological traits (days to flowering and maturity) (Table 5) indicated that season end moisture stress (mainly after flowering) might have caused relatively a yield reduction in the late maturing genotypes. Days to maturity was negatively correlated with all traits except with days to flowering (r = 0.56, P<0.05). Plant height had a weak and non-significant association with all traits. Number of capsules was highly and positively correlated with number of branches (r = 0.85, P<0.001), which is in line with several studies like (Yrgalem, *et al*., 2013; Akbar, 2011).

Table 5: Correlation coefficients between grain yield and yield traits of sesame genotypes

	YLD	DF	DM	LCBZ	NB	NC	PH	OC
YLD	_							
DF	-0.110^{ns}	_						
DM	-0.587*	0.566*	_					
LCBZ	0.332^{ns}	-0.477^{ns}	-0.607*	_				
NB	0.705**	0.425^{ns}	-0.274^{ns}	-0.122^{ns}	_			
NC	0.680**	0.266^{ns}	-0.377^{ns}	0.248^{ns}	0.847***	_		
PH	0.222^{ns}	0.282^{ns}	-0.184^{ns}	0.503^{ns}	0.144^{ns}	0.356^{ns}	_	
OC	0.798**	-0.442^{ns}	-0.544^{ns}	0.543^{ns}	0.073^{ns}	0.253^{ns}	0.251^{ns}	_

ns=non significant, *= significant at P<0.05, **=highly significant at P<0.01

It is obvious that sesame is grown mainly for its oil over all of the sesame growing areas. Correlation of an average oil content of the different genotypes and the BLUP (Best Linear Unbiased Predictor) values of other traits including yield were determined. Oil content was positively correlated with LCBZ (r = 0.543), NB (r = 0.073), NC (0.253), and PH (r = 0.251). However, it was significantly and positively correlated (r = 0.798) with grain yield only. Chowdhury, *et al*. (2010) were also reported a positive correlation between oil content and grain yield. Moreover, oil content was negatively correlated with DF (r = -0.442) and DM (r = -0.544) but neither of which were correlated significantly. As explained above, early maturing genotypes had better yield and also high oil content.

4. Conclusion

Based on the cluster analysis, the different genotypes were grouped in to five clusters and those genotypes under cluster-II (G1 and G4) were with highest mean grain yield. Hence, these genotypes might be better parental materials for further breeding program to improve sesame grain yield especially for the genotypes under cluster-

III, which are with greater genetic divergence (D^2=215.573).

Grain yield had a significant and positive correlation (P<0.01) with oil content (r = 0.79), number of branches (r = 0.7) and number of capsules (r =0.68). On the contrary, yield was negatively correlated with days to maturity (r = -0.58) and days to flowering (r = -0.11). The traits which showed positive association with grain yield had their own positive contribution in having better grain yield of the genotypes and those negatively correlated affected negatively the grain yield of the genotypes. As grain yield is a complex trait and is highly influenced by genetic as well as environmental factors, direct selection for yield may not be as such effective. Thus, exhaustive assessment of the association of grain yield and other agronomic traits is indispensably important for further breeding program and enhancing grain yield of sesame.

Acknowledgement

The first author would like to sincerely acknowledge for research members of crop department in Humera Agricultural Research Center and Public and Private Partnership Organization project (PPPO) for financial support.

Reference

Abd El- Mohsen, A. A. (2013). Comparison of some statistical techniques in evaluating Sesame yield and its contributing factors, *Scientia Agriculturae,*, 1 (1), pp.8-14.

Ahmed, M.B.S. and Ahmed, F.A. (2012). Genotype X season interaction and characters association of some Sesame (*Sesamum indicum* L.) genotypes under rain-fed conditions of Sudan, *African Journal of Plant Science,* 6(1), pp.39-42.

Akbar, F., Rabbani, M. A., Shinwari, K.Z and Khan, J.S. (2011). Genetic divergence in Sesame (*Sesamum indicum* L.) landraces based on qualitative and quantitative traits, *Pak. J. Bot.*, 43(6), pp. 2737-2744.

Anilakumar, R.K., Pal.A., Khanum, F. and Bawa, A.S. (2010). Nutritional, medicinal and industrial uses of Sesame (*Sesamum indicum* L.) Seeds, *Agric. conspec. sci*., 75(4).

Aremu, C.O., Adewale, D.B. and Adetunji, I. A. (2011). Cause and effect variations and trait selection ndex for indigenous Sesame (*Sesamum Indicum* L.) genotypes, *International Journal of Applied Agricultural and Apicultural Research,* 7 (2), pp. 64-71.

Ashri, A. (2007). Sesame (*Sesamum indium* L.). In: Singh, R.J., ed. *Genetic Resources, Chromosome Engineering, and Crop Improvement.* USA: CRC Press, Boca Raton, 4, pp. 231-289.

Chowdhury, S., Datta, A.K., Saha, A., Sengupta, S., Paul, R., Maity, S. and Das, A. (2010). Traits influencing yield in Sesame (*Sesamum indicum* L.) and multilocational trials of yield parameters in some desirable plant types, *Indian Journal of Science and Technology* , 3(2), pp.163-166.

Cornish, R. (2007). Statistics: Cluster Analysis.

FAOSTAT, (2012). (Food and Agriculture Organization of the United Nations). http://faostat.fao.org/

Kobayashi, T., Kinoshita, M., Hattori, S., Ogawa, T., Tsuboi, Y., Ishida, M., Ogawa, S. and Saito, H. (1990). Development of the Sesame metallic fuel performance code, *Nucl. Technol,* 89(2), pp.183-193.

Mahalanobis, P.C. (1936). On the generalized distances in statistics, *Proc. Nation: Inst. Sci.,India,* 2, pp. 49-55.

Minitab, (1998). MINITAB users guide, released 16.0, MINITAB Inc.

More, A. J., Bhoite, K. D. and Pardeshi, S.R. (2006), Genetic diversity studies in forage maize (*Zea mays* L.). Res. on Crops, 7(3): 728-730

Naturland, (2002). Organic farming in the tropics and subtropics. Producers guide.

Ong'Injo, E.O. and Ayiecho, P.O. (2009). Genotypic variability in Sesame mutant lines in Kenya, *African Crop Science Journal*, 17(2), pp. 101 - 107.

Parsaeian, M., Mirlohi, A. and Saeidi, G. (2011). Study of genetic variation in Sesame (*Sesamum indicum* L.) using agro-morphological traits, 47, pp.314-321.

Pham, D.T., Nguyen, T.T.D., Carlsson, A.S. and Bui, M.T. (2011). Morphological evaluation of Sesame (*Sesamum indicum* L.) varieties from different origins, *Australian Journal of Crop Science* 4(7), pp.498-504.

Piepho, H.P., MÖhring J., Melchinger A.E., Bùchse A. (2008). BLUP for phenotypic selection in plant breeding and variety testing, *Euphytica*, 161, pp. 209 – 228.

Searle S.R., Casella G., McCulloch, C.E. (1992). Variance components, Wiley, New York.

Singh, P., Sain, D., Dwivedi, V. K., Kumar, Y. and Sangwan, O. (2005). Genetic divergence studies in maize (*Zea mays* L.). Annals of Agri. Bio. Res., 10(1): 43-46.

Steel, R. and Torrie, J. (1980). Principles and Procedures of Statistics a Biometrical Approach. 2nd ed. Mc Graw-Hill, Inc, pp. 471-472.

Subramanian, S. and Subramanian, M. (1994). Correlation studies and path coefficient analysis in Sesame (*Sesamum indicum* L.), *J. Agron. Crop Sci*., 173,pp.241-248.

Tabatabaei, I., Pazouki, L., Bihamta, R.M., Mansoori, S., Javaran, J.M. and Niinemets, U. (2011). Genetic variation among Iranian Sesame (*Sesamum indicum* L.) accessions vis-à-vis exotic genotypes on the basis of morpho-physiological traits and RAPD markers, *AJCS*, 5(11), pp.1396-1407.

Tripathi, A., Bisen, R., Ahirwal, P., Paroha, S., Sahu, R. and Ranganatha, A. R. (2013). Study on genetic divergence in Sesame (*Sesamum indicum* L.) germplasm based on morphological and quality traits, *Supplement on Genetics & Plant Breeding,* 8(4), pp.1387-1391.

Urdan, T.C. (2005). *Statistics in Plain English*. 2nd ed. London, Mahwah, NewJersey: Lawrence Erlbaum Associates, Publishers.

Ward J.H. (1963). Hierarchical grouping to optimize an objective junction, *J. Am. Stat. Assoc*, 58, pp.236-244.

Yirgalem, T., sentayehu, A. and Geremew, T. (2013). Assessment of genetic variability, Genetic Advance, Correlation and Path analysis for Morphological Traits in Sesame Genotypes, *International Journal of Plant Breeding and Genetics*, 7(2), pp.21-34.

Yol, E., Karaman, E., Furat, S. and Uzun, B. (2010). Assessment of selection criteria in Sesame by using correlation coefficients, path and factor analyses, *Australian Journal of Crop Science*, 4(8), pp. 598-602.