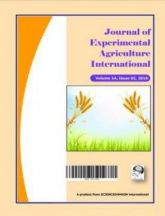

Journal of Experimental Agriculture International

27(3): 1-16, 2018; Article no.JEAI.43983
ISSN: 2457-0591
(Past name: American Journal of Experimental Agriculture, Past ISSN: 2231-0606)

Multivariate Analysis of Genetic Diversity in the Ethiopian Garden Cress (*Lepidium sativum* L.) Genotypes

Legesse Tadesse[1,2], Firew Mekbib[2], Adugna Wakjira[3] and Zerihun Tadele[4*]

[1]Department of Biology, College of Natural and Computational Sciences, Dire Dawa University, P.O.Box 1362, Dire Dawa, Ethiopia.
[2]School of Plant Sciences, College of Agriculture and Environmental Sciences, Haramaya University, P.O.Box 138, Dire Dawa, Ethiopia.
[3]Ethiopian Institute of Agricultural Research, P.O.Box 2003, Addis Ababa, Ethiopia.
[4]Institute of Plant Sciences, University of Bern, Altenbergrain 2, 3013 Bern, Switzerland.

Authors' contributions

This work was carried out in collaboration between all authors. Authors LT and FM designed the study. Author LT performed the experiment and statistical analysis, wrote the protocol and manuscript and revised the manuscript. Authors FM, AW and ZT supervised the experiment and revised both the protocol and manuscript and helped in data analysis. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/JEAI/2018/43983
Editor(s):
(1) Dr. Peter A. Roussos, Assistant Professor, Lab. Pomology, Agricultural University of Athens, Greece.
(2) Dr. Bishun Deo Prasad, Assistant Professor, Department of Molecular Biology & Genetic Engineering, Bihar Agriculture College, Sabour, Bihar Agricultural University, Bihar, India.
Reviewers:
(1) Nebi Bilir, Suleyman Demirel University, Turkey.
(2) Jamile da Silva Oliveira, Universidade de Brasília, Brazil.
Complete Peer review History: http://www.sciencedomain.org/review-history/26872

Original Research Article

Received 17 August 2018
Accepted 16 October 2018
Published 27 October 2018

ABSTRACT

Aims: The use of multivariate techniques is an important strategy for germplasm classification and study of genetic relationships among genotypes. This study was designed to evaluate using multivariate analysis the genetic divergence among 112 garden cress (*Lepidium sativum* L.) genotypes collected from different Administrative Zones of Ethiopia.
Methodology: The experiment was conducted at Haramaya University Research Site and Kulumsa farmer field. Twelve agro-morphological traits were evaluated in Randomized Complete Block Design (RCBD) with two replications. Region-wise analyses of variance, principal component and cluster analyses were applied.

**Corresponding author: E-mail: zerihun.tadele@ips.unibe.ch;*

Results: The analysis of diversity pattern based on the region of genotype origin revealed highly significant variability within and between regions of origin for almost all studied traits. The geographical pattern of distributions of genotypes in groups were not dependent on the regions of origin. The first three principal components explained 80.3% of the total variations suggesting that traits such as number of secondary branches, days to maturity, plant height, biomass/plant and biomass/plot, harvest index, grain yield/plant, thousand seed weight and grain yield/ha are the principal discriminatory traits in the germplasm studied. The cluster analysis categorised the 112 genotypes into six groups. The most diverse genotypes were found between cluster I and II which could be used for producing new genetic variability and exploitation of heterotic effects with the traits of interest in crossing programs.
Conclusion: The genetic diversity existing in the current study could be utilised in the genetic improvement of the Ethiopian garden cress germplasm.

Keywords: Cluster analysis; garden cress; genetic divergence; Lepidium sativum; multivariate analysis; principal component analysis.

1. INTRODUCTION

Garden cress (*Lepidium sativum* L.), locally known as *fetto* (Amharic) or *shimfu* (Afaan Oromo) [1], belongs to the family Brassicaceae (cabbage family) where it is known as a traditional herbal healer for a number of diseases [2,3,4] to the rural majority. A worldwide growing interest in the health benefits of herbs and botanicals has been increasing over years [5,6]. Currently garden cress has been given due attention for alleviating nutritional deficiencies and its medicinal properties [6]. Garden cress is among the neglected or under-utilised horticultural crops [3,7]. It is also an aromatic plant which contains a considerable amount of essential oil [8] and edible oil rich in its medicinal properties. Garden cress is used in the form of vegetable (sprout) in North Africa, West and Central Asia and United Kingdom besides its medicinal use. Vavilov [9] and recent publications [3,7] demonstrated that the highland region of Ethiopia and Eritrea are the primary center of origin and diversity for garden cress. However, until the present time, most of the genetic improvement programs on garden cress had not been well established in both the developed and developing countries. In Ethiopia, the crop is neglected by the research community hence received little research investment. Due to this, all garden cress accessions available at the national germplasm depository are landraces which are highly adapted to specific environmental conditions. Hence, there has been no improved variety cultivated in the country. The cultivation of these landraces has declined in recent years due to high disease incidence and poor productivity [10]. In addition, lack of adequate information on garden cress genotypes has led to inefficient use for both conservation and utilisation of its desirable traits.

For an efficient evaluation and utilisation of the genetic materials, detailed knowledge about genetic diversity, and information on collection and classification are important and the basis for crop improvement programs [11,12], which is elucidated through different marker systems such as agro-morphological, biochemical and molecular markers. Among these, agro-morphological characterisation is considered as the initial step [11,13] for designing breeding program. Temesgen et al. [14,15] and Said [10] on the study of 49 and 85 accessions, respectively reported the presence of broad based genetic variations of garden cress landraces in Ethiopia. It was known that there could be many genetic and environmental effects on the crop variation [16]. However, partly the number of genotypes involved, the test area and agronomic condition and the environments were not ideal and hence the genotypes were unable to release their full potential for characterisation. Thus, the studies were not inclusive, applying multivariate analysis and requiring comprehensive evaluation of the genetic divergence of garden cress.

For characterisation and classification of a large number of accessions, multivariate methods such as principal component and cluster analysis are considered to be the most appropriate for qualifying the degree of genetic diversity among the genotypes included in the study. These methods aid to evaluate not only the patterns of genetic variation but also the courses of developing target traits such as grain yield [14, 15,17,18]. Therefore, the present study was undertaken with the objectives to estimate the

magnitude of genetic divergence and identify the set of major contributing traits among regions of origin for further utilisation in breeding programs.

2. MATERIALS AND METHODS

2.1 Description of the Study Area

The field experiment was performed at Agricultural Research Site of Haramaya University (HRS) and at Kulumsa farmer field (KFF) during *Meher* cropping season of 2014/2015. HRS is located at 9°26'N and 42º3' E in Eastern Hararghe Zone at elevation of 2020 meter above sea level (m.a.s.l). The soil type of the location is fluvisols and its annual average rainfall is 786.8 mm. The mean minimum and maximum temperature are 8.3°C and 23.4°C, respectively with an average temperature of 16.8°C ([19]; Jigjiga Meteorological Station). KFF is located at 07º 59' 31" N and 39 º 11' 12" E in Arsi Zone of Oromia Regional State, and elevation of 2400 m.a.s.l. The site is situated 4.3 km east of Kulumsa Agricultural Research Center. It receives an average annual rainfall of 840 mm and the annual average temperature of 17ºC with maximum and a minimum temperature of 22.8ºC and 10.5ºC, respectively [20]. The soil type is Vertisols with a good drainage system. The area receives well distributed and extended rainfall both in amount and distribution creating relatively suitable environment for the cultivation of garden cress.

2.2 Experimental Procedures and Data Collection

The experimental materials used in this study are comprised of 112 genotypes of garden cress of which 90 were obtained from the Ethiopian Biodiversity Institute, while the remaining 22 landraces (CG1-CG22) were collected during the harvesting season (October 2013 to January 2014) from different agro-ecological regions of Ethiopia. The genotypes were grown in Randomized Complete Block Design (RCBD) with two replications. Each entry was sown in two rows of 135 cm length at spacing of 20 cm between rows, 15 cm between plants and 30 cm distance between plots. The seeds were drilled in the rows and thinned at 10 days after emergence till distance between plants within row became nearly 15 cm apart [3]. Fertiliser (DAP) application was performed on the basis of 100 kg per hectare during planting and all the recommended cultural practices, such as weeding and hoeing were applied uniformly in both experimental sites.

The selection and measurement of parameters were made by adopting the International Board for Plant Genetic Resources, Descriptors for Brassica and Raphanus [21] and Guide lines for Development of Descriptors List [22] in view specific descriptor for garden cress. The following data were collected on plant basis: Plant height (PH), Number of primary branches (PB), Number of secondary branches (SB), Grain yield/plant (GYPP) and Biomass/plant (BMPP) were recorded from five randomly sampled plants and their average were taken for each plot and location. On the other hand, plot basis data were collected from the entire plot on: days to flowering initiation (FI), days to 50 % flowering (FL), days to maturity (DM), thousand seed weight (TSW), grain yield/plot (GYPlt), harvest index (HI), biomass/plot (BMPlt), and grain yield/ha (GYPHa). Four genotypes collected from Amhara and Oromia Regional States were excluded from data collection because of their poor performance under field conditions (Table 1).

2.3 Data Analysis

A combined analysis of variance based on the genotypes pooled into their respective regions of origin [Amhara, Oromia, SNNP, Tigray, Somali and others] was carried out for the data generated from two locations. The homogeneity of residual variances and significant differences among the genotypes were conducted based on the standard statistical procedures [23,24]. Character means were standardised to mean zero and variance of unity to avoid differences in measurement scale [25,26] before undertaking principal component and cluster analysis. Principal component analysis was computed following the methods of Sneath and Sokal [27] to reduce the number of variables into a few uncorrelated components. Therefore, PCs with eigenvalues greater than unity were considered important in explaining much of the variability.

Genetic divergence was determined following Mahalanobis's [28] generalised distance (D^2). The Mahalanobis generalised distances were utilised to estimate the distances between and within clusters using the SAS computer software package as per the following formula: $D^2ij = (Xi-Xj)' S^{-1}(Xi-Xj)$. Where, D^2ij is the distance between any two groups i and j; Xi and Xj are the vector mean of the traits for the i^{th} and j^{th} groups

respectively, and S^{-1} is the inverse of the pooled covariance matrix. Average intra and inter cluster distances were calculated by the following formula as suggested by Singh and Chaudhary [29]: Average intra-cluster (D^2) $=\sum D^2/n$ and average inter-cluster distances (D^2) $=\sum D^2 i_1 j_2/ni_1 nj_2$. The contribution of each trait to divergence as described by Sharma [30] with the formula [(CD) = (SD/χ) X 100%] where SD, χ are the standard deviation and mean performance of each trait, respectively.

Cluster analysis based on Ward's method [31] using squared generalised distance metric and a dendrogram was constructed as described by Sneath and Sokal [27]. Duncan's New Multiple Range Test (DMRT) and/ or Least Significant Difference (LSD) was used to provide significant differences between averages of traits among clusters. The significance of D^2ij values for pairs of clusters were tested using the calculated values of chi-square (χ^2) at 1%, and 5% probability level. The test was done against the tabulated values of χ^2 for 'P' degrees of freedom, where P is the number of quantitative characters considered [29]. The PROC GLM, PROC PRINCOMP and PROC CLUSTER of SAS Version 9.1.3 [32], and Minitab version 17 [33] statistical softwares were used in the biplot analysis.

3. RESULTS AND DISCUSSION

3.1 Regional Patterns of Genetic Diversity

Wider ranges of variations were observed among garden cress genotypes for all investigated traits (Table 2). Pooled genotypes into region of origins depicted narrow range of variation compared to the entire ranges of all studied traits. Relatively higher range of coefficient of variation value (26.5 to 29%) were observed for grain yield per ha, number of secondary branches, biomass per plot and biomass per plant as well as grain yield per plant. The presence of high variations in these characters could also be evident from the wide range of values within the region than between the regions (Table 2). The wide variation in the mean grain yield per ha and yield related traits demonstrate the high potential for grain yield improvement by simple selection. Though both the entire and pooled genotypes showed significant variation in almost all measured traits, large amount of variations were noted within the region than between the regions of origins. Above all, number of secondary branches, biomass/plot and grain yield/ha showed high magnitude and relatively wider range of variation within and among regions.

Table 1. List of garden cress genotypes investigated and their origin

Origin (regional state)	Number of genotypes	Name of genotypes
Amhara	27	CG22, 207542, 241777, 229203, 235892, 229205, 229202, 229200, 205163, 215714, 208030, 229199, 90018, 229799, 212628, 90020, 229201, 229798, 229204, 215713, 90004, 205162, CG11, CG12, CG14, 214243[a], CG7[a]
Oromia	39	208769, 208693, 215808, 90006, 212852, 230831, 237991, 216885, 216816, 90022, 215807, 216886, 19001, 19000, 90021, 234828, 208666, 19002, 90005, 212853, 230830, 208669, 18841, 208667, 230524, 18843, CG1, CG2, CG3, CG5, CG6, CG8, CG9, CG10, CG15, CG19, CG20, 90002[a], CG4[a]
SNNP[b]	15	8604, 205141, 202116, 225725, 240396, 90016, 242916, 225799, 240578, 240808, 240397, CG13, CG16, CG17, CG18
Somali	4	230829, 230523, 216815, 231210
Tigray	17	219960, 238273, 219962, 219959, 233982, 234355, 233983, 233984, 219958, 233986, 237512, 233985, 219961, 242609, 207910, 233981, CG21
Others[c]	10	90012, 233679, 233370, 240579, 90010, 90009, 90014, 90017, 90007, 90008

[a]*Excluded from the final data analysis*
[b]*SNNP: Southern Nation Nationalities and People Regional States*
[c]*Originated in Ethiopia but the site of collection is not described*

Table 2. Pattern of genetic variability for genotypes and regions

Parameter[a]	For genotypes				For regions			
	Mean	SD	CV	Range	Mean	SD[c]	CV[d]	Range
FI	56.15	6.44	11	46.00 - 70.00	56.46	6.4	2.3	46.83 - 68.00
FL	61.45	6.91	11	51.00 - 74.00	61.8	7.5	3.8	51.33 - 73.17
PH	78.77	13.26	17	44.60 - 111.00	79	13.2	6.54	50.77-104.90
PB	17.70	3.25	18	10.00 - 44.00	18	3.4	11.4	10.94 - 28.74
SB	106.81	31.34	29	40.00 - 175.00	107	31.7	11.8	51.17 - 173.67
DM	114.79	10.05	9	101.00 - 130.00	115	11	2.2	101.00 - 130.00
BMPP	41.60	11.13	27	17.10 - 84.60	42	11	10.43	21.80 -74.18
BMPlt	390.03	103.8	27	181.80 - 732.00	390	103	2.35	216.60 - 659.18
GYPP	8.10	2.26	28	2.73 - 13.21	8	2.3	10.24	3.65 - 12.74
HI	19.77	3.84	19	8.35 - 31.77	20	3.6	10.94	10.65 - 28.41
TSW	2.34	0.33	14	1.25 - 3.42	2	0.3	8.47	1.61 - 3.13
GYPHa	2165.9	574.1	26.5	904.5[b] - 3484.8[b]	2160	575.7	7	1075.87 - 3407.78

[a]FI: days to flowering initiation; FL: days to 50% flowering; PH: plant height; PB: number of primary branches; SB: number of secondary branches; DM: days to maturity; BMPP: biomass/plant; BMPlt: biomass/plot; GYPP: grain yield/plant; HI: harvest index; TSW: thousand seed weight; GYPHa: grain yield/ha
[b]Recorded at HRS; [c]SD: standard deviation; [d]CV: coefficient of variation

Genotypes originated from Amhara and SNNP (Southern Nation Nationalities and People Regional States) recorded higher grain yield/ha (Table 3). As per the region wise analysis of variance, highly significant variation among the genotypes of the regions for the majority of the morphological traits is an indication of a high degree of variation implying the great potential of the genotypes in future breeding programs through selection. Similarly, a high degree of variation for days to flowering, days to maturity, plant height, pods per plant and seed yield per plant were reported for the Ethiopian lentil germplasm [34]. The results of this study were also partly in agreement with that of Said [10] which showed that genotypes from Amhara and Tigray Regional States had higher performance in grain yield than other Regional States. Genotypes from Somali Regional State scored relatively low grain yield/ha as they were lowland types and hence low yielding under the two testing sites of Kulumsa and Haramaya. This result was in agreement with that of Said [10].

Table 3. Region-wise means comparison for agro-morphological characters

Trait[a]	Regional states of germplasm origin						LSD (0.05)
	Amhara (25)[b]	Oromia (37)	SNNP (15)	Somali (4)	Tigray (17)	Others (10)	
FI	55.73c	55.83c	55.63c	56.69b	56.46b	58.43a	0.54
FL	61.32c	61.05c	61.07cb	62.38b	61.41b	63.48a	0.96
PH	79.04b	76.72c	78.76b	76.58c	79.53b	85.22a	2.19
PB	18.13a	17.19c	17.86c	17.08c	17.61b	18.71a	0.87
SB	111.09b	101.68d	105.26c	93.56e	109.32b	118.48a	4.36
DM	114.70b	114.27b	115.05b	114.13b	114.60b	117.15a	1.05
BMPP	42.90a	41.69b	40.20c	40.80c	39.47d	44.07a	1.76
BMPlt	401.55a	388.39b	386.84b	388.44b	373.49c	401.70a	4.46
GYPP	8.37a	8.01b	8.39a	7.92c	7.84c	8.10a	0.32
HI	19.61b	19.49c	20.93a	19.37c	20.16a	18.97c	0.89
TSW	2.357a	2.34a	2.30b	2.43a	2.29b	2.38a	0.08
GYPHa	2225.54a	2137.4b	2217.16a	2099.71b	2109.42b	2167.34a	64.8

[a]FI: days to flowering initiation; FL: days to 50% flowering; PH: plant height; PB: number of primary branches; SB: number of secondary branches; DM: days to maturity; BMPP: biomass/plant; BMPlt: biomass/plot; GYPP: grain yield/plant; HI: harvest index; TSW: thousand seed weight; GYPHa: grain yield/ha.
[b]Values in parenthesis represent the number of genotypes from the respective region. Means in a row followed by the same letter are not significantly different (P< 0.05) from each other.

The results of region wise analysis of variance indicated that the mean squares due to genotypes were highly significant for all characters studied (P< 0.001) (Table 4). The significance of variability due to origin of germplasm collection, environment, and genotype for the traits such as flowering initiation, 50% flowering, plant height, biomass/plot, harvest index, thousand seed weight, and grain yield/ha indicated the existence of plenty of genetic variation among the genotypes with differential performance across location. The results noted the presence of high potential and well performing genotypes for future yield improvement programs. Flowering initiation, biomass/plant, biomass/plot, harvest index, thousand seed weight, and grain yield/ha showed significant differential performance due to the interaction effects of region x location, and/or region x genotypes. The two environments showed relatively higher differential performance of biomass per plot and grain yield per ha (Table 4).

The high degree of variability and association between morphological characters and geographical origin showed genotypes in each region have significant variability for developing varieties having either wider or specific adaptability by exploiting the regional germplasm. The performance of genotypes due to the interaction of replication by location and replication by genotypes depicted none significant variations to the majority of the variables under study.

Plenty of variations were noted within the regions than between the regions of origins probably due to a free flow of genes across the region *via* exchange of seed stock between the regional states or directional selection by the growers and consumers (favoring some common features like seed colors or grain yield by the consumers regardless of origin of the crop). Similar findings were reported in Indian mustard [35], sorghum [36] and tetraploid wheat [37]. The adequate genetic variability of the studied genotypes is important for the improvement of the crop by considering both the phenotypic performance of individual genotypes as well as sources of origin of genotypes.

3.2 Principal Component Analysis

The analysis indicated the first three PCs, each with eigenvalues greater than unity explained 80.3% of the total variation among the studied genotypes for all morphological characters (Table 5). The total variations in the 12 measured variables of the genotypes accounted 52.5% for PC1, 17.9% for PC 2 and, 10% for PC 3. The first principal component accounted for maximum variability (52.5%) in the data with respect to succeeding components and hence measurement of independent impact of plant height, number of secondary branches, days to maturity, biomass/plant, biomass/plot, and grain yield/plant and grain yield/ha to the total variation as described by Syafii et al. [12]. While, grain yield/plant, thousand seed weight, harvest index and grain yield/ha were the major positive contributors, days to flowering initiation, days to 50% flowering accounted for negative variability in PC2. The variation in PC3 mainly was attributed to harvest index with positive loading value and biomass/plant with negative loading value (Table 5). The higher coefficients (eigenvector) for some traits designated the relatedness of that character to relevant PC axes [27].

The presence of positive and negative correlations between the components and the variables are interpreted by positive and negative loading values. Characters with the largest absolute value within the first PC are influencing the cluster more than those with lower absolute value closer to zero [38]. The extent of variance and relationship among different variables as demonstrated by loading plot indicated the extent of association among measured variables (Fig. 1A). The loading plot of PC 1 against PC 2 shows the relationships among variables based on pooled performances of the landraces across the two locations. Similar relationships were observed among the variables for the biplot analysis where the position of each genotypes were plotted (Fig. 1B). While all variables had positive values for the first component, they were roughly equally divided into positive and negative values for the second component. The importance and relationship between variables within a component are determined by the magnitude and direction of a factor loadings within a PC [39]. The sign of the loading indicates the direction of the relationship between the components and the variable. The greater the loading factors the higher contribution of the associated traits to the variance. There were relatively strong association between FI and FL; DM, PB, and SB; BMPP and BMPlt; GYPP and GYPHa; while, TSW and HI had relatively weak association between themselves and with the rest of traits.

Table 4. Mean square values for agro-morphological characters of garden cress based on source of origin

Trait[a]	Sources of variations[b]						Residual (df=215)	R^{2c} (%)	CV (%)[d]
	Region (df=5)	Location (df=1)	Genotype (df=107)	L x G (df=107)	R x L (df=5)	R x G (df=85)			
FI	53.34***	165.02***	147.05***	12.00***	5.020*	2.38*	1.64	98.00	2.28
FL	42.30***	58.52**	163.41***	14.84***	11.05ns	5.95ns	5.340	97.00	3.77
PH	481.51***	5464.18***	524.85***	52.36***	39.84ns	24.06ns	27.39	96.00	6.64
PB	21.24***	7.95ns	23.10***	9.21***	6.18ns	5.13ns	4.13	89.00	11.45
SB	2912.2***	61001.6***	2344.7***	711.78**	127.61ns	87.9ns	103.30	95.60	9.5
DM	55.42***	4947.79***	315.49***	28.07***	4.54ns	9.11ns	6.15	98.00	2.20
BMPP	170.69***	8735.31***	305.53***	35.87***	1.02ns	55.09***	20.81	95.80	10.96
BMPlt	7896.7***	1216180.8***	28159.6***	3196.52***	520.54***	168.98*	108.7	99.70	2.67
GYPP	2.97***	828.62***	9.25***	2.42***	1.2ns	0.35ns	0.62	96.50	9.7
HI	26.62***	608.00***	26.78***	10.75***	4.06ns	11.54***	4.78	91.13	10.94
TSW	0.09*	6.47***	0.18***	0.10***	0.03ns	0.05*	0.04	90.00	8.47
GYPHa	184132.98***	36805577.7***	751923.34***	178303.75***	88399.63**	16841.96ns	24158.9	98.00	7.20

[a]*FI: days to flowering initiation; FL: days to 50% flowering; PH: plant height; PB: number of primary branches; SB: number of secondary branches; DM: days to maturity; BMPP: biomass/plant; BMPlt: biomass/plot; GYPP: grain yield/plant; HI: harvest index; TSW: thousand seed weight; GYPHa: grain yield/ha*

[b]*L x G: location x genotype; R x L: region x location; R x G: region x genotype*

[c]*Coefficient of determination*

[d]*Coefficient of variation*

, **, *: Significant at 5%, 1% and 0.1 % probability levels, respectively*

Table 5. Eigenvalue, proportion of variance and traits that contributed to first three principal components

Trait	PC1	PC2	PC3
Flowering initiation (days)	0.287	-0.396	0.104
Days to 50% flowering (days)	0.281	-0.400	0.071
Plant height (cm)	0.324	-0.201	0.133
Primary branches (No.)	0.258	-0.106	-0.038
Secondary branches(No.)	0.348	-0.015	0.026
Days to maturity (days)	0.318	-0.250	0.130
Biomass/plant (g)	0.315	0.150	-0.478
Biomass/plot (g)	0.325	0.161	-0.150
Grain yield/plant (g)	0.326	0.345	0.075
Harvest index	0.049	0.308	0.804
1000-seed weight (g)	0.147	0.436	-0.190
Grain yield/ha (kg/ha)	0.331	0.340	0.089
Eigenvalue	6.298	2.145	1.193
Proportion of variance	0.525	0.179	0.099
Cumulative total variation	0.525	0.704	0.803

The major contributing traits in PC1, PC2 and PC3 have played the most important role, allowing the breeders to improve yield of the crop as well as save a lot of time for the identification and selection of agronomically better performing genotypes. Similar findings were reported by different authors on the study of different brassica species [35,40,41]. The discriminatory power was evenly distributed across eleven traits therefore differentiating group of genotypes mainly attributed by collective effect of individual traits (Table 5) and hence challenge for selection of parent plants linked to these minor effects of traits for breeding programs. In this study, almost all evaluated traits showed significantly higher contributors for 80.3% of total variation. Sabaghnia et al. [18] also reported the first four factors, with eigenvalues greater than unity, explained about 80% of the total variation in garden cress. Similar results of major contributing traits for the total variation were reported for different crops [12,35,41] for yield and yield related traits.

3.3 Cluster Analysis

Success of any crop improvement programs mainly depends on amount of diversity available in the crop. The results of cluster analysis for 12 studied traits in 108 genotypes are presented in Table 5. The cluster analysis was used to obtain a dendrogram of garden cress genotypes (Fig. 2) with different mean values of the traits. In D^2 distance based dendrogram for agronomic traits, all the 108 genotypes were distributed based on their relatedness into six main clusters. It is interesting to note that genotypes from Amhara and Oromia Regional States were represented in all six clusters. This might be due to at least two reasons, i) the 59% of germplasm included in the current study were collected from the two regions, and ii) the two regions represent the largest geographical area in the country, hence expected to hold huge diversity in garden cress populations. The geographical pattern of diversity in cluster analysis indicated that the genotypes belonging to diverse ecological regions clustered together, while those of same region entered separate groups. Since differences in agro-morphological parameters of genotypes did not follow their geographic pattern of origin, genotypes from each region had its own distinctness and were grouped independently. For instance, genotypes from Somali region clustered into two groups, while cluster IV contained genotypes from all regions (Table 6). These results were in agreement with the findings of Temesgen et al. [15] and Khan et al. [11].

The possible reason for grouping of genotypes from different origin of collection in a single cluster could be the exchange of germplasm among garden cress growers from diverse regions, or unidirectional selection practiced by farmers in tailoring the promising landraces for different regions. As suggested earlier by Endashaw [42] based on the study on tef (*Eragrostis tef*), genotypes in different regions narrowed into similar genetic materials as a result of genetic drift, natural variation and artificial selection other than ecological and geographical diversifications. On the other hand, studies by Singh et al. [43] on *Coriandrum*

sativum L and by Khan et al. [11] on *Brassica napus* L., respectively did not show consistent relationships between genetic divergence and geographical pattern of diversity. Hence, the diversification of garden cress genotypes in different agro-ecologies of Ethiopia observed in current study might be due to these divergence and convergence distribution causing factors.

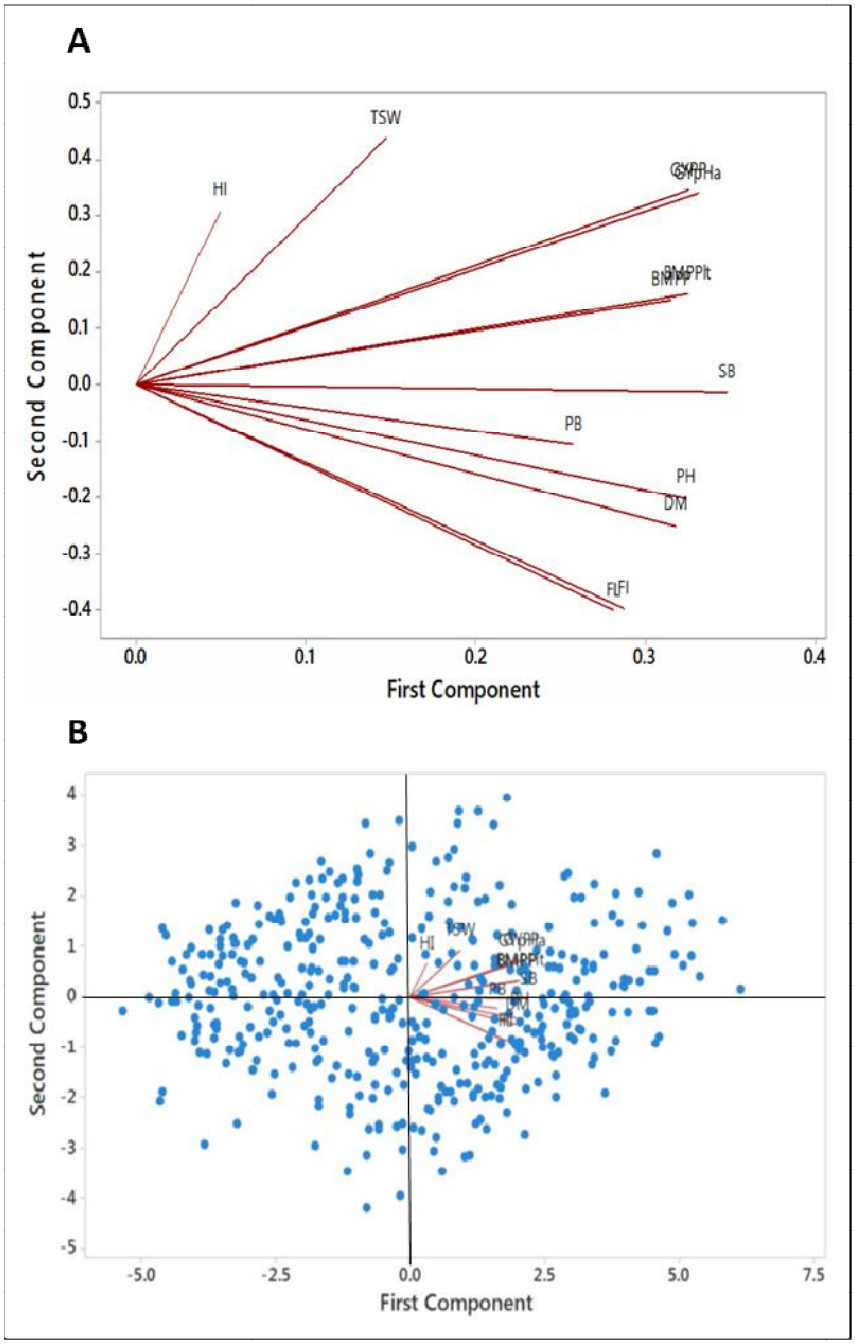

Fig. 1. Association of traits and genotypes. (A) Loading plot, and (B) Biplot analysis. FI: days to flowering initiation; FL: days to 50% flowering; PH: plant height; PB: number of primary branches; SB: number of secondary branches; DM: days to maturity; BMPP: biomass/plant; BMPlt: biomass/plot; GYPP: grain yield/plant; HI: harvest index; TSW: thousand seed weight; GYPHa: grain yield/ha

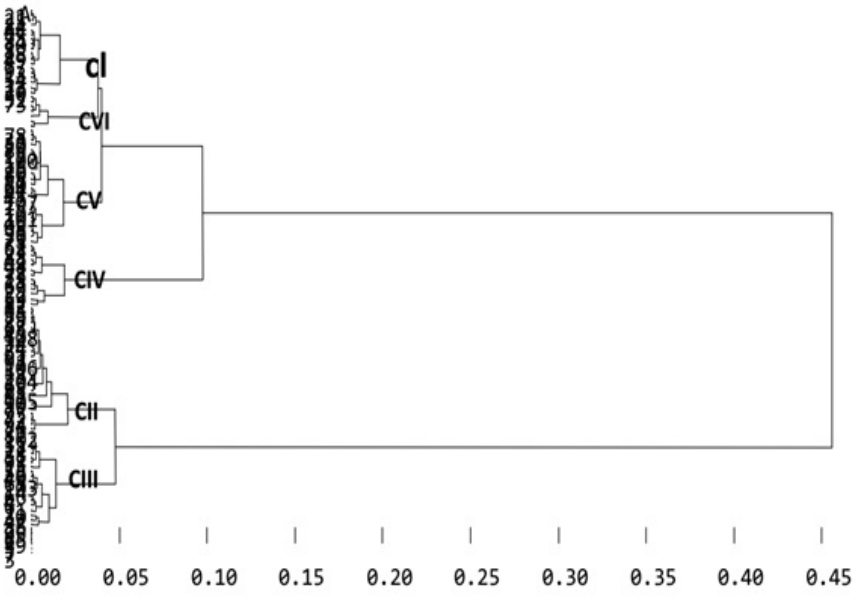

Fig. 2. Dendrogram depicting genetic relationships among genotypes using Wards linkage methods. Roman numbers indicate cluster numbers. Names of genotypes under each cluster are shown in Table 6

Table 6. Distribution of garden cress genotypes in different clusters

Cluster (number)	Name of genotype	Origin[a]
I (18)	219960, 238273, 219962, 219959, 237991, 235892, 233982, 234355, 233983, 229202, 219958, 215807, 208030, 234828, 242609, 212853, 229204, 90004	1, 2, 5
II (26)	225725, 216885, 230829, 242916, 225799, 230523, 240578, 208666, 212628, 90005, 240397, 231210, 230830, 208669, 90008, 230524, CG2, CG3, CG5, CG6, CG8, CG9, CG16, CG18, CG19, CG20	1, 2, 3, 4
III (20)	208769, 208693, 8604, 205141, 202116, 215808, 90006, 212852, 230831, 90016, 240579, 205163, 19000, 240808, 90010, 90014, 19002, 205162, CG13, CG17	1, 2, 3, 6
IV (13)	240396, 233679, 233370, 90022, 233985, 90018, 207910, 216815, 18841, 208667, 215713, 233981, CG22	1, 2, 3, 4, 5, 6
V (24)	207542, 241777, 90012, 229205, 233984, 233986, 229200, 15714,237512, 216886, 19001, 219961, 229199, 229799, 90020, 90009, 229201, 90007, 18843, CG10, CG11, CG12, CG15, CG21	1, 2, 5, 6
VI (7)	229203, 216816, 90021, 229798, 90017, CG1, CG14	1, 2, 6

[a]*Regional states: 1: Amhara; 2: Oromia; 3: SNNP; 4: Somali; 5: Tigray; 6: unknown source*

The late maturing (124 days) genotypes with the tallest plant height (89.74 cm) of 18 genotypes were grouped into cluster I (Table 6). They produced intermediate grain yield per hectare (2125.2 kg/ha) next to cluster V (2490.1 kg/ha). In contrast, cluster II comprised 26 genotypes, characterised with early maturity (103 days), relatively better yielder (2031.1 kg/ha) than cluster III (1614.3 kg/ha). Cluster 3 had 20 genotypes with early maturing (111 days) and intermediate thousand seed weight (2.26 gm), but with the lowest grain yield/ha (1614.30 kg/ha).

Cluster IV is the group with the highest performing genotypes, hence can be considered as a base population in the yield improvement programs. It has comprised 13 genotypes and was characterised by tall plant height (87 cm), late maturing (120 days), higher number of primary branches (20.3) and secondary branches (132.6) with corresponding higher biomass/plants

(57.5) and biomass/plots (503.7) and higher grain yield/ plant (10.4). Furthermore, this group contained genotypes with peculiar attributes of the greatest performance in grain yield/ha (2830.5 kg/ha) followed by genotypes in cluster V. Cluster V contained 24 genotypes with better performance in most yield and yield related traits. This cluster was characterised by long plant height, late maturing, higher grain yield/plant and thousand seed weight and high grain yield/ha (2490.1 kg/ha). Low performing and early maturing seven genotypes were grouped into cluster VI. The group contains genotypes with relatively long plant height (79.4 cm), late days to flowering initiation (60.3 days) and intermediate days to maturity (118.61 days) associated with low grain yield/ha (2001 kg/ha) next to Cluster III (Table 6).

Number of secondary branches, grain yield/plant, biomass/plant, biomass/plot, and grain yield/ha were major contributors of genetic divergence in the entire genotypes (Table 7) as described by Sharma [30]. The remaining traits had intermediate or very little contribution towards genetic divergence and hence, they were of less importance in assessment of genetic diversity. Similar kinds of observations were made by earlier workers [44]. Since varieties with narrow genetic base are increasingly vulnerable to diseases and adverse climatic changes, availability of the genetically diverse genotypes for hybridisation program become important.

The higher divergence in the grain yield and yield related traits could give a high opportunity for the development of high yielding cultivars through selection in the segregating generation. In line with this, Singh et al. [35] reported in the study of Indian mustard that, the variability in seed yield potential among different clusters could be used for exploitation of heterotic effects. The results obtained in the present study were supported in at least three earlier studies involving garden cress. [15], *Brassica napus* L. [11] and *Brassica carinata* L. [45].

Generally, a wide range of variability was observed among the 108 genotypes of Ethiopian garden cress with respect to seed yield per ha. The seed yield ranged from 1298.9 kg/ha to 3124.8 kg/ha with the gap of 1826 kg/ha between the lowest and the highest values indicating huge variability among populations in the current study. The overall mean yield from the current study was closer to the one reported for garden cress genotypes [14]. However, the mean seed yield reported in the current study (2420 kg/ha) was almost double of the one earlier reported for the same crop [18]. The lowest yield of the latter might be due to biotic and abiotic stresses. These pointed out that there could be greater potential of improving the grain yield *via* intensive improvement programs. Therefore, comprehensive collection focusing on the desired traits will benefit breeders at large for effective improvement in garden cress genotypes of Ethiopian collection.

3.4 Inter and Intra-cluster Divergence

Inter and intra cluster divergence values between and within six clusters were presented in the Table 8. The magnitudes of inter-cluster distance (D^2) were generally high and were indicators for the presence of substantial genetic diversity in Ethiopian garden cress. The intra-cluster distance is lower than the inter-cluster distances, implying that the accessions included within a cluster had less diversity among themselves. The average D^2 values ranged from 2.9 to 59.5. Intra cluster distances exhibited a range from 2.9 to 5.5 and inter cluster distance ranged from 11.3 to 59.5 (Table 8). The intra cluster distance was the maximum for Cluster VI (5.5) and minimum for Cluster II (2.9). Cluster VI with seven genotypes, exhibited intra-cluster distance value (5.5) and genotypes from this cluster could be utilised as parental lines for hybrid breeding program or recombination breeding program owing to their wider within group distance.

In general, cluster VI contained the most divergent genotypes within the group. The magnitudes of divergence between cluster I and Cluster II was maximum (59.5), while it was lowest (11.3) between Cluster I and V. Number of secondary branches, biomass per plant and grain yield/ha, biomass per plot, and grain yield per plant contributed for more than 59% of the total divergence at inter cluster level (Table 7). In agreement with this findings, number of primary branches and number of secondary branches were shown to contribute for the highest divergence in Indian mustard (*Brassica juncea*) [44]. As Pali and Mehta [46] confirmed in their study of flax selection of parents should be based on both the regional diversity and the magnitude of genetic divergence of clusters that make more economical and effective to screen divergent parents. Based on these, it could be concluded that 18 genotypes in Cluster I and 26 genotypes in Cluster II can be utilised for successful hybridisation program (Fig. 2, Table 7).

Table 7. The mean values of each trait investigated for the six clusters

CL	FI	FL	PH	PB	SB	DM	BMPP	BMPlt	GYPP	HI	TSW	GYpHa
I	63.32 a	68.51a	89.74 a	19.20b	125.95b	123.53a	44.54b	423.00c	8.00c	18.06c	2.14d	2125.20c
II	48.98 f	53.42 f	65.61e	16.1 d	82.31d	102.81f	37.04d	341.70d	7.65c	20.82b	2.39b	2031.10d
III	52.10e	57.68e	69.36d	15.21e	84.05d	110.18e	34.74d	317.09d	6.14e	17.77c	2.26c	1614.30e
IV	58.46c	64.15c	86.92b	20.32a	131.58a	119.88c	57.51a	503.70a	10.40a	18.61c	2.67a	2830.50a
V	60.27b	66.35b	88.01ab	19.15b	124.88b	121.20b	43.71b	430.20b	9.27b	21.46ab	2.38b	2490.10b
VI	57.50d	62.00d	79.44c	17.14c	105.71c	118.61d	33.84d	344.30d	7.44c	22.35a	2.11d	2001.00d
Mean	56.77	62.02	79.85	17.85	109.08	116.04	41.90	393.33	8.15	19.85	2.32	2182.03
LSD	1.3	1.4	3.05	0.97	8.39	2.19	3.11	30.37	0.67	1.26	0.11	158.39
SD	5.31	5.63	10.27	2.01	21.89	7.92	8.87	71.14	1.49	1.94	0.20	423.44
CIC	9.4	9.1	12.9	11.3	20.1	6.8	21.2	18.1	18.3	9.8	8.6	19.4
CGL	0.11	0.11	0.17	0.18	0.29	0.09	0.27	0.27	0.28	0.19	0.14	0.27

Abbreviations: CIC: contribution of traits for inter cluster divergence; CGL: contribution of traits for divergence at genotypic level; CL: cluster; FI: days to flowering initiation; FL: days to 50% flowering; PH: plant height; PB: number of primary branches; SB: number of secondary branches; DM: days to maturity; BMPP: biomass/plant; BMPlt: biomass/plot; GYPP: grain yield/plant; HI: harvest index; TSW: thousand seed weight; GYPHa: grain yield/ha

Means followed by the same letter are not significantly different ($p< 0.05$); "a" is the highest mean of the variable

Table 8. Intra-cluster (bolded diagonal) and inter-cluster (off diagonal) distance (D^2) values across genotypes

Cluster	I	II	III	IV	V	VI
I	**3.58**					
II	59.50**	**2.85**				
III	39.75**	15.71	**3.37**			
IV	31.46*	45.76**	53.50**	**4.23**		
V	11.26	49.64**	43.90**	20.27*	**3.01**	
VI	14.33	35.39 **	24.51*	39.80**	16.20	**5.47**

, significant at 5% and 1% level of significance; X^2 = 18.31 and 23.21 at 5%, 1% probability level respectively*

The crosses between divergent genotypes (Clusters I and II) produce higher genetic variability, maximum heterotic performance and hence maximising the chance of improving traits of interest. The superior derivatives may not be expected from crosses between genotypes included in Clusters I and II due to low mean grain yield performance of both clusters. Cluster I and V exhibited minimum genetic distance (11.3) between them, which showed that genotypes in these two clusters were somewhat similar in genetic constitution and hybridisation between these groups may not generate sufficient variability.

Therefore, based on cluster distances of the present study, genotypes in Clusters IV and V can be utilised as donor parent for improving productivity due to their high mean performance for grain yield/ha and most of the yield contributing traits with high amount of genetic divergence. As Tsehaye and Kebebew [47] and Saha et al. [48] verified, the greater the distance between the two clusters, the wider the genetic diversity among the parents for exploiting their heterotic effect in hybridisation program. Hence, genotypes of garden cress with wide genetic variation accompanied with useful characteristics effectively employed in intra specific crosses with the hope that this would lead to the transmission of higher genetic gain for different yield related and quality traits. In short, genotypes clustered in distant groups could be used as parents for hybridisation to exploit heterosis for improving yield, early maturity and resistance to biotic and abiotic stresses.

4. CONCLUSION

The current study showed a high level of variations in traits investigated among genotypes and within the regional states of Ethiopia. Variability within regions was higher than that of between regions. Principal component (PC) analysis has identified major variables accounting for 80.3% of the total variations. The first three PCs included diverse agro-morphological traits (such as biomass/plot, grain yield/plot, harvest index, and grain yield/ha) that play a prominent role in classifying the germplasm collections and accelerating the selection procedure. The total genotypes were clustered into six groups, and the clusters with higher inter-cluster divergence (Cluster I and II) can be used as a parental material for further breeding to produce desirable recombinants. The crosses from these divergent genotypes are expected to manifest maximum heterosis and wide variability in genetic architecture. Clustering of genotypes did not follow the pattern of geographic distribution. In general, the results of this investigation indicated that substantial variability present in the garden cress population will provide an opportunity for breeders to exploit the diversity through hybridising the most divergent genotypes followed by selection.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Amare Getahun. Some common medicinal and poisonous plants used in Ethiopian folk medicine. Faculty of science, Addis Ababa University, Addis Ababa, Ethiopia; 1976.
2. Sharma S, Agarwal N. Nourishing and healing powers of garden cress (*Lepidium sativum* L.): A review Indian Journal of Natural Products and Resources. 2011;2: 292-297.

3. Wadhwa MS, Panwar A, Agrawal N, Saini I, Patidar LN. A review on pharmacognostical study of *Lepidium sativum*. Advanced Research in Pharmaceuticals and Biologicals. 2012;2: 316-323.
4. Doke S, Guha M. Garden cress (*Lepidium sativum* L.) Seed an important medicinal sources. Journal of Natural Products of Plant Resources. 2014;4:69-80.
5. Mali GR, Mahajan GS, Mehta AA. *Lepidium sativum* (Garden cress): A review of contemporary literature and medicinal properties. Oriental Pharmacy and Experimental Medicine. 2007;7:331-335.
6. Tahseen MA, Mishra G. Ethnobotany and diuretic activity of some selected Indian medicinal plants: A scientific review. The Pharma Innovation Journal. 2013;2:109-121.
7. Bermejo H, León. Neglected Crops: From a Different Perspective. J. E. Plant Production and Protection Series No. 26. FAO, Rome, Italy; 1994.
8. Nigist A, Sebsebe D. Aromatic Plants of Ethiopia. Shama Books, Addis Ababa, Ethiopia; 2009.
9. Vavilov NI. Phytogeographic Basis of Plant Breeding. The origin, variation, immunity and breeding of cultivated plants. Chronica Botanica. New York, Ronald Press. 1951; 13:1-366.
10. Said M. Genetic diversity study of *Lepidium sativum* populations from Ethiopia using morphological characters and ISSR markers. MSc. Thesis, Addis Ababa University, Addis Abeba, Ethiopia; 2012.
11. Khan SA, Iqbal J, Khurshid H, Saleem N, Rabbani MA, Zia M, Shinwari ZK. The extent of intra-specific genetic divergence in *Brassica napus* L. population estimated through various agro-morphological traits. European Academic Research. 2014;2: 2255-2275.
12. Syafii M, Cartika I Ruswandi D. Multivariate analysis of genetic diversity among some maize genotypes under maize-*albizia* cropping system in Indonesia. Asian Journal of Crop Science. 2015;7:244-255.
13. Smith JSC, Smith OS. The description and assessment of distance between inbred lines of maize. The utility of morphological, biochemical and genetic descriptors and a scheme for the testing of distinctiveness between inbred lines. Maydica. 1989;34: 151-161.
14. Temesgen B, Mebeaselassie A, Million E. Genetic variability and association among yield, yield related traits and oil content in Ethiopian garden cress (*Lepidium sativum* L.) genotypes. Journal of Plant Breeding and Crop Sciences. 2013a;5(7):141–149.
15. Temesgen B, Mebeaselassie A, Million E. Genetic divergence analysis of garden cress (*Lepidium sativum* L.). Journal of Plant Breeding and Crop Sciences. 2013b; 5:770–774.
16. Yazici N, Bilir N. Aspectual fertility variation and its effect on gene diversity of seeds in natural stands of *Taurus Cedar* (*Cedrus libani* A. Rich.). International Journal of Genomics. 2017;2017:2960624.
17. Hegay S, Geleta M, Bryngelsson T, Asanaliev A, Garkava G, Hovmalm HP, Ortiz R. Genetic diversity analysis in *Phaseolus vulgaris* L. using morphological traits. Genetic Resources and Crop Evolution. 2014;61:555–566.
18. Sabaghnia N, Ahadnezhad A, Janmohammdi M. Genetic variation in garden cress (*Lepidium sativum* L.) germplasm as assessed by some morphological traits. Genetic Resources and Crop Evolution. 2015;62:733–745.
19. Wassu M, Solomon A, Beneberu S, Simret B. Genetic Diversity of Local and Introduced Sweet Potato [*Ipomoea batatas* (L.) Lam.] Collections for Agro Physicochemical Attributes in Ethiopia. STAR Journal. 2015;4:9-19.
20. Gezahegn F, Sentayehu A, Zerihun T. Genetic variability studies in bread wheat (*Triticum aestivum* L.) genotypes at Kulumsa Agricultural Research Center, South East Ethiopia. Journal of Biology, Agriculture and Healthcare. 2015;5(7):89-98.
21. IPBGR. Descriptors for Brassica and Raphanus. International Board for Plant Genetic Resources, Italy Rome; 1990.
22. Bioversity International. Guidelines for the development of crop descriptor lists. Bioversity Technical Bulletin Series. Bioversity International, Rome, Italy; 2007.
23. Gomez KA, Gomez AA. Statistical Procedure for Agricultural Research. 2nd Edition. J. Wiley and Sons, New York; 1984.

24. Steel GD, Torrie JH. Principles and Procedures of Statistics: A Biometrical Approach. 2nd edn, McGraw Hill Book Co., New York, USA; 1980.
25. Rao CR. Advanced statistical methods in Biometric Research. John Wiley & Sons, Inc., New York; 1952.
26. Saed-Moucheshi A, Fasihfar E, Hasheminasab H, Rahmani A, Ahmadi A. A review on applied multivariate statistical techniques in agriculture and plant science. International Journal of Agronomy and Plant Production. 2013;4(1):127-141.
27. Sneath PHA, Sokal RR. Numerical taxonomy; the principles and practice of numerical classification. W.H. Freeman and Co., San Francisco, USA; 1973.
28. Mahalanobis PC. On the generalized distance in statistics. Proceedings of National Institute of Science. 1936;2:49-55.
29. Singh RK, Chaudhary BD. Biometrical Methods in Quantitative Genetic Analysis. Kalyani publishers, New Delhi-Ludhiana, India; 1985.
30. Sharma JR. Statistical and biometrical techniques in plant breeding, New age internationals limited, publisher, New Delhi Banglore, India; 1998.
31. Ward JR. Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association. 1963; 58(301):236-244.
32. SAS Institute INC. SAS/STAT user's guide. Version 9. Fourth Edition. Statistical Analysis Institute Inc., Cary North Carolina; 2004.
33. Minitab. Version 17 Statistical Software by Minitab Inc. Harrisburg; 2010.
34. Edossa F, Kassahun T, Endashaw B. A comparative study of morphological and molecular diversity in Ethiopian lentil (*Lens culinaris* Medik.) landraces. African Journal of Plant Sciences. 2010;4:242-254.
35. Singh KH, Shakya R, Mahawar RK. Genetic diversity and patterns of variation among Indian mustard (*Brassica Juncea* (L.) Czernj. and Cosson) Varieties SABRAO. Journal of Breeding and Genetics. 2014;46(2):329-339.
36. Geleta N, Labuschagne MT. Qualitative traits variation in sorghum (*Sorghum bicolor* (L.) Moench) germplasm from eastern highlands of Ethiopia. Biodiversity and Conservation. 2005;14:3055-3064.
37. Hailu F, Merker A, Johnson E. Patterns of phenotypic diversity for phenologic and qualitative traits in Ethiopian durum wheat germplasm. Genetic Resources and Crop Evolution. 2010;53:1089-1098.
38. Chahal GS, Gosal, SS. Principles and Procedures of Plant Breeding Biotechnology and Conventional Approaches. Narosa Publishing House, New Delhi; 2002.
39. Azeez MA, Aremu CO, Olaniyan OO. Assessment of genetic variation in accessions of sesame (*Sesamum indicum* L.) and its crosses by seed protein electrophoresis. J. Agroaliment. Process. Technol. 2013;19:383-391.
40. Yousuf M, Ajmal SU, Munir M, Ghafoor A. Genetic diversity analysis for agro-morphological and seed quality traits in rapeseed (*Brassica campestris* L.). Pakistan Journal of Botany. 2011;43(2): 1195-1203.
41. Yared S, Misteru T. Variability, heritability and genetic advance analysis for some morphological characters in oilseed brassica breeding lines. Molecular Plant Breeding. 2016;7(20):1-8.
42. Endashaw B. Morphological analysis of *Eragrostis tef*: detection for regional variation. Sinet: Ethiopian Journal of Sciences. 1996;19:117–140.
43. Singh SP, Katiyar RS, Rai SK, Tripayhi SM, Srivastava JP. Genetic divergence and its implication in breeding of desired plant type in coriander (*Coriandrum sativum* L.). Genetika. 2005;37(2):155-163.
44. Lodhi B, Thakral NK, Singh D, Avtar R, Bahadur R. Genetic diversity analysis in Indian mustard (*Brassica juncea*). Journal of Oilseed Brassica. 2013;4(2):57-60.
45. Zada M, Zakir N, Rabbani MA, Shinwari ZK. Assessment of genetic variation in Ethiopian mustard (*Brassica carinata* A. Brun) germplasm using multivariate techniques. Pakistan Journal of Botany. 2013;45(1):583-593.
46. Pali V Mehta N. Evaluation of genetic divergence in Indian flax (*Linum usitatissimum* L.). International Quarterly Journals of Life Sciences. 2015;10(4): 2043-2047.
47. Tsehaye T, Kebebew F. Morphological diversity and geographic distribution of adaptive traits in finger millet [*Eleusine coracana* (L.) Gaertn. (Poaceae)]

populations from Ethiopia. Ethiopian Journal of Biological Sciences. 2002;1:37-62.

48. Saha S, Begum T, Dasgupta T. Analysis of Genotypic Diversity in Sesame Based on Morphological and Agronomic Traits. Proceedings of the Conference on International Research on Food Security, September 19-21, 2012, Gottingen, Germany; 2012.

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sciencedomain.org/review-history/26872