Multivariate analysis in parental lines and land races of pearl millet \textit{[Pennisetum glaucum (L.) R. Br.]} 

Venkataramana. S. Kalagare\textsuperscript{1*}, N. Meenakshi ganesan\textsuperscript{1}, K. Iyanar\textsuperscript{2}, T. Chitdeshwari\textsuperscript{3} and C. N. Chandrasekhar\textsuperscript{4}  

\textsuperscript{1}Department of Genetics and plant Breeding, Tamil Nadu Agricultural University, Coimbatore  
\textsuperscript{2}Department of Millets, Tamil Nadu Agricultural University, Coimbatore  
\textsuperscript{3}Department of Soil Science & Agricultural Chemistry, Tamil Nadu Agricultural University, Coimbatore  
\textsuperscript{4}Department of Crop Physiology, Tamil Nadu Agricultural University, Coimbatore  
\textsuperscript{*E-Mail: vkalagare@gmail.com} 

Abstract 

Present investigation was carried out with 31 pearl millet genotypes containing R lines, B lines and land races at Department of Millets, Tamil Nadu Agricultural University, Coimbatore during kharif 2019 in order to assess the genetic diversity and to compare different methods of multivariate analysis. Mahalanobis’ $D^2$ cluster analysis resulted in six clusters with the highest inter cluster distance observed between cluster V and cluster VI. Cluster mean showed that cluster III and cluster I genotypes played significant influence on yield and yield attributing characters. In agglomerative hierarchical cluster (AHC) analysis, the highest inter cluster distance was observed between cluster II and cluster VI. Cluster V and cluster I represented the highest cluster mean for yield and yield component traits. Principal component analysis (PCA) resulted four principal components with eigen values more than one explaining 73.2 per cent variability. The biplot revealed five clusters with cluster I am possessing maximum number of genotypes and positively associated to most of the traits. PT 6706, PT 6709, Nattu Cumbu, Cumbu 2, PT 6676 and PT 6067 were the top-ranking genotypes upon PCA analysis with positive PC1 scores. All three multivariate analyses revealed considerable divergence in the experimental material as well as a comparable type of clustering in the diversity of R line, B line and small seeded land races and hence can be used in future breeding programmes.

Keywords: Mahalanobis’ $D^2$, Agglomerative hierarchical clustering, Principal component analysis, Pearl millet

INTRODUCTION 

Pearl millet \textit{[Pennisetum glaucum (L.) R. Br.]} is a small-grained tropical $C_4$ cereal crop grown in most adverse agroclimatic conditions \citep{Kumar2020} and important in the lives of the poor and low-income groups \citep{Govindaraj2020}. Pearl millet has a global area of approximately 340 lakh hectares and a yield of 310 lakh tonnes, ranking sixth among grains after wheat, rice, maize, barley, and sorghum \citep{Patil2020}. India is the largest producer of pearl millet in the world, with a production of 97.0 lakh tonnes from 75 lakh ha area \citep{INDIATESTAT}. In India, pearl millet hybrids account for 70\% of total pearl millet land area, with the remainder occupied by open-pollinated varieties (OPVs) or landraces \citep{Patil2020} indicating the importance of hybrid development in India. Pearl millet was grown on around 0.63 lakh hectares in Tamil Nadu, with productivity of 2,277 kg/ ha \citep{StatisticalHandbook}. Understanding the genetics and diversity of pearl millet will help in opening the door to additional possibilities for using it as a fodder and grain...
crop in today’s variable environment. Broadening the genetic base is crucial for boosting genetic gain for yield by introducing diverse germplasm accessions to attain maximum heterosis. Hybrid development programme is strongly depending upon the selection of diverse seed and pollen parent to develop high yielding hybrids (Sharma et al., 2020). Multivariate analysis, such as cluster analysis and principal component analysis (PCA) are the statistical procedures used to create the cluster in order to classify and identify divergent parents.

The $D^2$ statistics proposed by Mahalanobis is the most appropriate method for selecting morphologically divergent parents as it furnishes a measure of actual variation between any pair of populations (Malik et al., 2017). Mahalanobis’s generalized distance is estimated by $D^2$ statistics for discriminating population considering a set of parameters together rather than inferring from indices based on morphological similarities and polygenic relationship (Sankar et al., 2014; Singh and Gupta, 1979; Rasitha et al., 2020; Swamynatham et al., 2020). Hierarchical cluster analysis is a commonly used method for forming clusters and displaying similarities and dissimilarities between pairs of genotypes, in which agglomerative hierarchical clustering were formed by grouping cases into bigger and bigger clusters until all cases are members of a single cluster. Principal components analysis (PCA) is the data reduction technique applicable to quantitative type of data and transforms multi-correlated variables into another set of uncorrelated variables (Kumar et al., 2020).

The current study employed multivariate analysis, comprising cluster analysis (Agglomerative hierarchical clustering (AHC) and Mahalanobis’ $D^2$ statistics) and principal component analysis (PCA) with the objective to examine the diversity of 31 important pearl millet lines and to make comparison of the different methods.

**MATERIALS AND METHODS**

The current investigation was carried out at the Department of Millets, Tamil Nadu Agricultural University, Coimbatore which lies in western agroclimatic zone of Tamil Nadu, India. This zone has an altitude in the range of 200 m to 600 m and located between 11°55’ to 10°02’ N latitude and 76°51 to 78°09’ latitude. Field experiment was conducted during the main cropping season kharif (June to October) during 2019 at Department of Millets, Tamil Nadu Agricultural University, Coimbatore which is located at 11°01’ to 30.7°N longitude and 76°55’ to 35.0°E latitude. Total of 31 pearl millet genotypes which included 17 restores lines, 3 maintainer lines, 10 land races and an open pollinated variety Dhanashakti was utilized for the present investigation. Experimental material was laid out in randomized complete block design (RCBD) with two replications. A total of 16 quantitative traits were recorded on randomly selected five competitive plants in two replications except days to 50% spike emergence, which was a single observation by visual assessment of group of plants in each replication on plot basis. Then, the mean data in each replication were subjected to statistical analysis.

The statistical analysis of replicated data was carried out with the help of the software WINDOSTAT ver 7.1 for $D^2$ statistics. $D^2$ statistics was originally developed by Mahalanobis (1936) and Rao (1952) suggested its application for the assessment of genetic diversity in plant breeding. The genotypes were grouped on the basis of minimum generalized distance using the Tocher’s methods. Software XLSTAT was used for employing AHC (Agglomerative Hierarchical Clustering) method, where Euclidean distance between the genotypes were calculated from the unweighted pair group method using arithmetic averages (UPGMA) and PCA (Principal Component Analysis) was carried out by software XLSTAT Version 2014.5.0 for standardized mean data. Cluster diagram for AHC was analysed through software Graphical Genotypes (GGT 2.0).

**RESULTS AND DISCUSSION**

Mahalanobis’s $D^2$ statistics is a technique for the assessment of genetic diversity in various breeding materials. For exploiting heterosis in hybrid development programme, it is necessary to utilize parents with maximum genetic divergence. More diverse the parents, more are the chances of pronounced heterotic effects and increased spectrum of variability in the segregating generations (Govindaraj et al., 2011). Mahalanobis’s $D^2$ analysis employed for the grouping 31 pearl millet genotypes using 16 yield and yield attributing traits resulted in six major clusters. The dendrogram for the $D^2$ cluster analysis using Tocher method is depicted in the Fig. 1. Wilk’s Criterion simultaneous test of significance showed that there was highly significant difference among genotypes for all the characters. Out of six clusters, cluster I possessed maximum genotypes of 23 followed by four genotypes in cluster II. Most of R (restorer line) and B (maintainer line) lines were in cluster I and all small seeded land races like Kuttu cumbu 1, Kuttu cumbu 2 and Kuttu cumbu 3 were in cluster II. Pattern of distinct clusters and allotment of land races and breeding lines in different clusters indicated the presence of divergence between land races and lines used in experimental material. Out of three B lines, two B lines (ICMB 98222, ICMB 99222) came under cluster I and cluster VII possessed one B line (ICMB 06111). This indicated the clear differentiation between maintainer lines (Kaushik et al., 2018). Among the land races, the genotypes Cumbu 1, Shoolgiri local and Nattucumbu were fell in cluster I, whereas cluster III possessed Kizikuppam local and cluster II accounted for four small seeded land races. These uneven distribution of land races to different clusters and most lines falling into few clusters suggested that land races collected from the same geographic area were not necessarily closely related and different regions did not necessarily
have different genetic background (Upadhyay and Murty, 1970; Dave and Joshi, 1995 and Govindaraj et al., 2011). Genetic drift and selection under different environments could have caused greater divergence than geographical distance (Upadhyay and Murty, 1970).

Higher inter-cluster distance was observed than intra-cluster distance (Table 1). The maximum intra cluster distance was observed in cluster I followed by cluster II and least in cluster III, IV, V and VI as these had only one genotype. The maximum inter cluster distance was observed between cluster V and cluster VI followed by cluster III and cluster VI, cluster IV and cluster VI and minimum cluster distance was observed between cluster IV and cluster V. The higher inter-cluster distance than intra-cluster distance showed homogeneity and narrow genetic variability within a cluster (Singh and Gupta, 1979; Malik et al., 2017). Lesser intra cluster distance indicated that the genotypes inside the cluster should be nearly identical in their characteristics and less divergent. More intra cluster distance may be due to degree of heterogeneity and pedigree and hence, selection will be efficient if it is based on highest mean for desirable traits (Ramya et al., 2017; Kaushik et al., 2018; Rasitha et al., 2020). The genotypes within the clusters can be subjected to further analysis for morphological trait uniformity and to test for general combining ability to combine 4 to 10 lines to develop synthetics and composites. Maximum inter cluster distance was observed for cluster V, cluster VI and cluster III, which had PT6583, ICMB06111 and Kizikuppam local genotypes, respectively. These genotypes could be utilized for evaluation of specific combining ability (sca), hybrid development and to obtain good recombinants in F$_2$ (Govindaraj et al., 2011; Sankar et al., 2014; Ramya et al., 2017; Rasitha et al., 2020). Maximum inter cluster distance also lead to wide spectrum of variability in segregating population to operate selection (Singh et al., 1981; Govindaraj et al., 2011; Athoni et al., 2016). It was also indicated that the genotypes with the restricted genetic divergence having relatively smaller statistical distances or falling in the same cluster were also likely to produce desirable heterotic effects in the population resulting from crossing if they complement some major weaknesses of each other as against those involving genotypes which falling in distant clusters and possessing wide genetic divergence.

Cluster II showed highest mean for the number of productive tillers. Cluster III recorded highest mean for leaf length, flag leaf length, flag leaf width, thousand seed weight and single plant yield. Cluster IV recorded maximum mean for the traits like leaf width, number of nodes, panicle diameter total, number of grains per panicle and biological yield. Cluster V recorded highest mean for leaf sheath length, flag leaf width, number of nodes, panicle length and plant height. Cluster VI showed maximum mean for the trait days to flowering and harvest index (Table 2). From the results of cluster mean values, it was clear that for good yield and yield attributing traits, cluster III played prominent role and it included land race Kizikuppam local followed by the cluster I which included almost all-important R and B lines. Cluster II comprised of small seeded land races with high cluster mean for number of productive tillers and plant height with fair amount of biological yield could be used for the fodder purpose (Dave and Joshi, 1995). It was also important to note that, the cluster mean for days to 50% spike emergence was low in cluster III and cluster II, indicating early maturity enabling the cultivars to escape from terminal drought (Govindaraj et al., 2011; Sumathi, et al., 2016; Rasitha et al., 2020).

Highest percent contribution to the total variability was due to thousand grain weight followed by single plant yield, plant height, panicle length, days to 50% spike emergence and number of productive tillers (Table 2). Nearly 90 per cent of variability was contributed by thousand grain weight, single plant yield, plant height, number of nodes, biological yield, leaf width, panicle length, days to 50% spike emergence and panicle diameter indicating the opportunity of these traits for selection in the given experimental material

Table 1. Estimates of intra (diagonal bolded) and inter (non-diagonal non-bolded) cluster distances in pearl millet genotypes for yield and yield attributing traits by D$_2$ method

| Clusters | Cluster I | Cluster II | Cluster III | Cluster IV | Cluster V | Cluster VI |
|----------|-----------|------------|-------------|------------|-----------|------------|
| Cluster I | 11.37     | (3.37)     | 13.22       | 13.99      | 13.91     | 17.03      |
| Cluster II | 16.54     | (4.07)     | 15.93       | 15.22      | 15.56     | 18.73      |
| Cluster III | 10.71     | (3.27)     | 19.53       | (3.9)      | (3.94)    | (4.33)     |
| Cluster IV | 0         | (0)        | 19.96       | 19.29      | 20.11     | 20.11      |
| Cluster V  | 0         | (0)        | 19.47       | 19.39      | 20.07     | 20.89      |
| Cluster VI | 13.22     | (4.47)     | 19.29       | (4.39)     | (4.48)    | (4.57)     |

Values in the parenthesis are D$^2$ distance

https://doi.org/10.37992/2022.1301.023
Fig. 1. Dendrogram representing the D2cluster analysis using Tocher method in pearl millet genotypes for yield and yield attributing traits
Table 2. Cluster mean of pearl millet genotypes and percentage contribution to total variability for yield and yield attributing traits by D2 method

| Cluster | DTF | LSL | LL | LW | FL | FW | NN | PL | PD | NT | PH | TGP | TSW | BY | SPY | HI |
|---------|-----|-----|----|----|----|----|----|----|----|----|----|-----|-----|----|-----|----|
| Cluster I | 50.39 | 17.84 | 57.25 | 4.27 | 44.16 | 4.25 | 6.34 | 23.48 | 3.06 | 3.23 | 157.03 | 2739.81 | 10.92 | 75.94 | 51.22 | 70.91 |
| Cluster II | 40.25 | 12.19 | 45.13 | 3.09 | 29.13 | 3.26 | 5.92 | 18.50 | 1.72 | 4.83 | 173.66 | 1506.95 | 4.80 | 77.33 | 30.63 | 41.86 |
| Cluster III | 40.00 | 16.75 | 73.00 | 3.95 | 53.00 | 4.50 | 5.34 | 25.33 | 2.92 | 3.66 | 170.54 | 2486.64 | 12.20 | 71.50 | 79.00 | 110.84 |
| Cluster IV | 50.00 | 19.00 | 64.00 | 4.76 | 39.67 | 4.23 | 7.00 | 31.33 | 3.15 | 3.33 | 160.00 | 4144.96 | 6.10 | 109.64 | 43.12 | 42.90 |
| Cluster V | 52.00 | 23.33 | 48.33 | 4.20 | 43.33 | 4.50 | 7.00 | 31.67 | 3.04 | 2.66 | 177.91 | 1944.42 | 9.92 | 99.77 | 31.58 | 31.85 |
| Cluster VI | 58.00 | 14.66 | 34.00 | 3.46 | 25.33 | 3.50 | 4.00 | 15.66 | 2.68 | 3.33 | 83.67 | 2304.33 | 7.05 | 32.74 | 43.45 | 132.76 |
| Per cent contribution | 4.3 | 1.94 | 0.22 | 7.1 | 1.08 | 1.29 | 8.39 | 4.95 | 3.87 | 1.29 | 9.46 | 3.01 | 28.6 | 7.74 | 12.9 | 3.87 |

DTF- Days to 50% spike emergence, LSL- Leaf sheath length (cm), LL- Leaf length (cm), LW- Leaf width (cm), FL- Flag leaf length (cm), FW- Flag leaf width (cm), NN- Number of nodes, PL- Panicle length (cm), PD- Panicle diameter (cm), NT- Number of productive tillers, PH- Plant height (cm), TGP- Total number of grains per panicle, TSW- Thousand grain weight (g), BY- Biological yield (g), SPY- Single plant yield (g) and HI- Harvest index (%)
Fig. 2. Dendrogram depicting pearl millet genotypes derived by AHC method of clustering for yield and yield attributing traits

Table 3. Estimates of intra and inter cluster analysis of 31 pearl millet genotypes for yield and yield attributing traits by AHC method

| Cluster | Cluster I | Cluster II | Cluster III | Cluster IV | Cluster V | Cluster VI |
|---------|-----------|------------|-------------|------------|-----------|------------|
| Cluster I | 19.15     | 21.25      | 9.93        | 19.85      | 13.03     | 34.24      |
| Cluster II | 22.83     | 20.44      | 28.77       | 24.88      | 40.26     |            |
| Cluster III | 0.00      | 0.00       | 20.60       | 34.77      |           |            |
| Cluster IV |           |            | 28.45       | 15.74      | 34.77     | 31.78      |
| Cluster V  |           |            |             | 6.76       | 11.25     |            |
| Cluster VI |           |            |             |           |           |            |
attributing traits indicating the significance of these lines to become potential parents for yield contributing traits. Out of three B lines, two B lines (ICMB 98222 and ICMB 99222) and out of 17 R lines three R lines (PT 6675, PT6674 and PT 6029) fell under cluster II, which showed the highest mean for panicle diameter and total number of grains per panicle. These lines could be used as potential parents for obtaining the panicles of bigger size with a greater number of seeds per panicle. The highest mean for the number of productive tillers and plant height observed for cluster VI included small seeded land races which could be used for forage purpose. The genotypes with contrast mean performance from these clusters could be utilized as potential parents in the development of hybrids for harnessing heterosis (Drabo et al., 2020). PC with higher eigen values and correlation was used to study interrelationship between different characters. Principal component analysis with correlation matrix is best to determine the principal factors, as it does not require the normal distribution assumption of populations (Chaudhary et al., 2015; Nehra et al., 2017; Santos et al., 2017; Sharma et al., 2020).

Principal component analysis is an effective approach for reducing the variability in multiple characters to the principal components with the first principal component capturing the maximum variability. The PCA based on correlation was used to study interrelationship between different characters. Principal component analysis with correlation matrix is best to determine the principal factors, as it does not require the normal distribution assumption of populations (Chaudhary et al., 2015; Sharma et al., 2020). PC with higher eigen values and variables with high factor loadings were considered as best representative of system attributes. In the present investigation, PCA was performed for yield and other attributes in pearl millet genotypes (Table 5 and Fig. 3). The first four principal components accounted for 73.20% of the total variability with eigen values more than one. The PC1 accounted for 37.55% of total variability followed by PC2, PC3 and PC4 exhibited 17.93%, 10.59% and 7.13% of total variability respectively. From the result it was revealed that maximum variability was spread within first four principal components where PC1 showed the highest variability among four. Hence, it was recommended to consider the characters or genotypes lying near and showing more PC1 score for catching the variability of particular trait (Ramya et al., 2017; Jain and Diwan, 2021). From the factor loading of PCA analysis, it was revealed that PC1 accounted maximum variability for most of the traits and other traits such as number of productive tillers, plant height and biological yield were captured in PC2. The harvest index and single plant yield were accounted in PC3 (Table 6). Result of factor loading of PCA analysis indicated that the maximum variability accounted by the PC1 was highly related to most of the yield attributing traits. PC2 showed maximum factor loading for number of productive tillers, plant height and biological yield, which were related to small seeded land races. Since, PC3 captured maximum variability for single plant yield and harvest index, the genotypes captured under this component can be utilized in improvement of crop for above mentioned traits. Characters which showed high positive or high negative contributed more to the diversity. The sign here indicates the relationship between variable and principal components. PC1 showed negative factor loading for number of productive tillers (-0.31) indicating the negative correlation with the trait (Ghazy et al., 2015; Chaudhary et al., 2015; Malik et al., 2017; Rasitha et al., 2020; Ramya et al., 2017; Jain and Diwan, 2021).

PCA results were generally are displayed as a biplot, in which axes correspond to the new system of coordinates (Fig. 4). The direction of arrow denotes the maximum change in great quantity and the length could be related with the rate of change occur. The acute coordinate angle (<90°) between the traits or principal component axis and trait shows the positive association between these traits, whereas obtuse angle (>90°) shows negative and right angle (=90°) indicates no correlation between the traits (Govindaraj et al., 2020). Most of the traits were in acute angle with the PC1 coordinates except number of productive tillers. The third quadrant did not have
Table 5. Eigen values and estimates of per cent variability accounted by the principal component analysis

| PC1  | PC2  | PC3  | PC4  | PC5  | PC6  | PC7  | PC8  | PC9  | PC10 | PC11 | PC12 | PC13 | PC14 | PC15 | PC16 |
|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| 6.01 | 2.87 | 1.69 | 1.14 | 0.95 | 0.86 | 0.65 | 0.57 | 0.39 | 0.31 | 0.21 | 0.12 | 0.09 | 0.08 | 0.05 | 0.01 |
| 37.55| 17.93| 10.59| 7.13 | 5.91 | 5.35 | 4.05 | 2.41 | 1.92 | 1.30 | 0.76 | 0.59 | 0.51 | 0.30 | 0.09 |
| 37.55| 55.49| 66.07| 73.20| 79.12| 84.47| 88.52| 92.12| 94.53| 96.46| 97.76| 98.51| 99.10| 99.61| 99.91| 100.00|

Fig. 3. Scree plot showing eigen values and percentage of cumulative variability

Table 6. Factor loading of four important principal components of pearl millet genotypes for yield and yield attributing traits

| S. No. | Traits      | PC1   | PC2   | PC3   | PC4   |
|--------|-------------|-------|-------|-------|-------|
| 1      | DTF         | 0.55  | -0.47 | -0.09 | -0.11 |
| 2      | LSL         | 0.77  | -0.11 | -0.21 | -0.09 |
| 3      | LL          | 0.74  | 0.33  | 0.06  | -0.08 |
| 4      | LW          | 0.84  | -0.14 | -0.19 | -0.18 |
| 5      | FL          | 0.86  | 0.13  | -0.01 | -0.23 |
| 6      | FW          | 0.73  | -0.01 | -0.10 | -0.38 |
| 7      | NN          | 0.41  | 0.40  | -0.13 | 0.34  |
| 8      | PL          | 0.71  | 0.20  | -0.33 | -0.20 |
| 9      | PD          | 0.71  | -0.44 | 0.05  | 0.44  |
| 10     | NT          | -0.31 | 0.75  | 0.31  | 0.04  |
| 11     | PH          | 0.28  | 0.84  | -0.06 | -0.11 |
| 12     | TGP         | 0.52  | -0.45 | 0.05  | 0.46  |
| 13     | TSW         | 0.69  | -0.06 | 0.38  | 0.27  |
| 14     | BY          | 0.34  | 0.65  | -0.25 | 0.36  |
| 15     | SPY         | 0.57  | 0.37  | 0.68  | 0.11  |
| 16     | HI          | 0.22  | -0.23 | 0.84  | -0.31 |

DTF - Days to 50 % spike emergence , LSL - Leaf sheath length , LL - leaf length , LW - Leaf width , FL - Flag leaf length, FW - Flag leaf width , NN - Number of nodes , PL - Panicle length , PD - Panicle diameter, NT - Number of productive tillers , PH - plant height, TGP – Total number of grains per panicle, TSW - Thousand grain weight , BY - Biological yield , SPY - Single plant yield and HI - Harvest index

https://doi.org/10.37992/2022.1301.023
any traits. All most all the traits were in acute angle with single plant yield. The lowest acute and adjacent angle with single plant yield were observed for leaf length and number of nodes. Number of productive tillers showed obtuse angle, which indicated the negative correlation between the traits. Single plant yield showed acute angle with most of the traits indicating positive correlation and exhibited the significance of the trait selection for improvement of yield attributing characters mainly like leaf length, number of nodes, panicle length etc. which were highly correlated (Kalagare et al., 2021; Sharma et al., 2020; Pujar et al., 2020).

Five different clusters were observed, where cluster I with the maximum number of genotypes possessed most of the genotypes on first quadrant followed by cluster II with 6 genotypes resided on third quadrant. Cluster I possessed most of the R lines and no B lines, whereas cluster II showed the two B lines (ICMB 98222 and ICMB 99222) and cluster IV showed one B line (ICMB 06111) on second quadrant. Cluster III possessed small seeded land races (Kuttu Cumbu 1, Kuttu Cumbu 2, Kuttu Cumbu 3 and Pothu Cumbu) on second quadrant. Biplot representing the genotypes is highly beneficial to identify the genotypes and character association and making the cluster for identification of diverse parents for utilization in development of hybrid. On the biplot, genotypes which closure to each other are similar and farthest are divergent (Sharma et al., 2020). The distance between the location of any two genotypes on the score plot is indirectly proportional to the degree of similarity. Genotypes which are nearer to the origin are contributing less to the variability, while those far from the origin are extremes and mostly extremes are favourable for breeding programme (Kiprotich et al., 2015). Similar genotypes formed group on the biplot which were differentiated by clusters. Results clearly indicated that the cluster formed on PC could differentiate the R lines, B lines and small seeded land races. Clusters contributing to the variability depended on how much far away the cluster formed from the origin. Cluster I and cluster II were nearer to the origin, whereas cluster III, cluster IV and cluster V were somewhat far away from the origin compared to cluster I and cluster II. Land races belonging to the cluster III contributed more to the variability, while those far from the origin indicating their stability and less variation (Chaudhary et al., 2015; Rasitha et al., 2020; Sharma et al., 2020).

The relationship between yield attributing traits with genotypes on first two principal components presented in the Fig.4. Among the five-clusters observed, cluster I with maximum genotypes was positively correlated with most of the traits except number of productive tillers. Most of the genotypes in the clusters I were bold seeded and possessed high single plant yield. Cluster III possessed small seeded land races and showed positive correlation with number of productive tillers, plant height and biological yield and negative correlation with days to flowering, harvest index, total number of grains per panicle and panicle diameter. The genotypes, PT 6067, Dnanashakti, PT 6581, PT6580, Uthankarai local and Shoolagiri local were nearer to the origin. The genotypes, PT 6707, PT 6067, Dnanashakti, PT 6581, PT6580, PT 6705, Uthankarai local and Shoolagiri local were nearer to the origin indicating their stability and less variation for the characters. Cluster III with small seeded four land races showed positive relationship with number of productive tillers, plant height and biological yield and can be involved in the breeding for fodder crops. Cluster I was almost opposite the cluster II and cluster IV indicated the diversity between the clusters. Accessions from diverse group will maximize opportunities to obtain transgressive segregants as there is a higher chance from genotypes to contribute unique desirable alleles at various loci. Hence, it is recommended to use the genotypes present in cluster I, cluster II, cluster IV and cluster V to intercross among these clusters (Chaudhary et al., 2015; Rasitha et al., 2020; Sharma et al., 2020).

List of top ten pearl millet accessions based on their PC score were arranged in a descending order of their scores (Table 7). The genotypes PT 6706, PT 6709, Nattu Cumbu, PT 6580, Cumbu 1, Cumbu 2, PT 6676, PT 6708, Kizikuppam local and PT 6067 showed maximum scores in PC1. The PC2 captured maximum score for the genotypes as Kuttu Cumbu 1, Pothu Cumbu, Cumbu 2, Kuttu Cumbu 2, PT 6580, Cumbu 1, Kizikuppam local, Kuttu Cumbu 3, PT 6076 and Dhanashakti. PC3 possessed maximum score for the genotypes like Kizikuppam local, ICMB 06111, Nattu Cumbu, Dhanashakti, Cumbu 2, PT 6705, PT 6582, PT 6581, PT 6706 and PT 6675. In present investigation, genotypes were identified through PC scores, where top ten PC1 scores were obtained by PT 6706, PT 6709, Nattu Cumbu, PT 6580, Cumbu 1, Cumbu 2, PT 6676, PT 6708, Kizikuppam local and PT 6067 showed maximum scores in PC1. The PC2 captured maximum score for the genotypes as Kuttu Cumbu 1, Pothu Cumbu, Cumbu 2, Kuttu Cumbu 2, PT 6580, Cumbu 1, Kizikuppam local, Kuttu Cumbu 3, PT 6076 and Dhanashakti. These results were in concordance with the biplot determination (Fig. 4) (Chaudhary et al., 2015; Sharma et al., 2020). The above small seeded land races showed earliness, which is an important heat escaping attributes of pearl millet and could be used as potential parents in breeding programmes for early flowering as explained by Malik et al. (2017). The factor loading for single plant yield and harvest index were high for PC2 and the genotypes Kizikuppam local, ICMB 06111, Nattu Cumbu, Dhanashakti. Cumbu 2 were the top five scoring genotypes in PC2 indicating the expected significant yield improvement by these genotypes. All three multivariate analysis (D2 statistics, AHC
Fig. 4. Biplot representing the relationship between yield and yield attributing traits of pearl millet genotypes on first two principal components

Table 7. List of top ten pearl millet accessions based on their PC scores

| Rank | PC1    | PC2    | PC3       |
|------|--------|--------|-----------|
|      | Genotypes | Score  | Genotypes | Score  | Genotypes | Score  |
| 1    | PT 6706  | 3.73   | KuttuCumbu1 | 3.89   | Kizikuppam local | 2.43   |
| 2    | PT 6709  | 2.59   | PothuCumbu | 2.06   | ICMB 06111 | 2.34   |
| 3    | NattuCumbu | 2.52  | Cumbu 2   | 2.01   | NattuCumbu | 1.98   |
| 4    | PT 6580  | 2.32   | KuttuCumbu2 | 1.80   | Dhanashakti | 1.95   |
| 5    | Cumbu 1  | 2.16   | PT 6580   | 1.79   | Cumbu 2   | 1.70   |
| 6    | Cumbu 2  | 1.99   | Cumbu 1   | 1.31   | PT 6705   | 1.16   |
| 7    | PT 6676  | 1.85   | Kizikuppam local | 1.19   | PT 6582   | 1.11   |
| 8    | PT 6708  | 1.49   | KuttuCumbu3 | 0.88   | PT 6581   | 1.08   |
| 9    | Kizikuppam local | 1.44 | PT 6706 | 0.71   | PT 6706   | 0.93   |
| 10   | PT 6067  | 1.43   | Dhanashakti | 0.70   | PT 6675   | 0.68   |

DTF - Days to 50% spike emergence, LSL - Leaf sheath length, LL - Leaf length, LW - Leaf width, FL - Flag leaf length, FW - Flag leaf width, NN - Number of nodes, PL - Panicle length, PD - Panicle diameter, NT - Number of productive tillers, PH - Plant height, TGP - Total number of grains per panicle, TSW - Thousand grain weight, BY - Biological yield, SPY - Single plant yield, and HI - Harvest index.
Table 8. Identification of diverse parents and land races through different multivariate techniques

| S. No. | D² statistics | Agglomerative hierarchical clustering | Principal component analysis |
|--------|---------------|---------------------------------------|-----------------------------|
| 1      | Kuttu cumbu 1, Kuttu cumbu 2, Kuttu cumbu 3 and Pothu Cumbu | Kuttu cumbu 1, Kuttu cumbu 2, Kuttu cumbu 3 and Pothu Cumbu | Kuttu cumbu 1, Kuttu cumbu 2, Kuttu cumbu 3 and Pothu Cumbu |
| 2      | Kizikuppam local | Kizikuppam local Cumbu 2               | Kizikuppam local Cumbu 2    |
| 3      | PT 6676         | PT 6059                               | PT 6059                     |
| 4      | PT 6583         | PT 6674, PT 6675, ICMB 98222, ICMB 99222, PT 6029 | PT 6674, PT 6675, ICMB 98222, ICMB 99222, PT 6029, PT 6710 |
| 5      | ICMB 06111      | ICMB 06111                             | ICMB 06111                  |

method and PCA) showed that there was a significant divergence between 31 genotypes. The cluster analysis Mahalanobis' D² statistics and Agglomerative hierarchical clustering (AHC) showed the variation in number of clusters and grouping of genotypes into different clusters. This is mainly because of the method of clustering and the type of data input for the cluster analysis. In case of Mahalanobis’ D² statistics replicated data were used as input. During the testing of the significance (Wilks’ Criterion simultaneous test of significance), it excluded the replication variances and consider only the determinants of error and error plus varietal variances. For transformation of the original variable to uncorrelated variable by pivotal condensation method and for calculation of D² values it uses error variance and covariance. The cluster formation by Tocher method was used where the cluster formation depends on the arbitrary value of D² which is generally approximately near to the maximum D² value between any two population (Singh and Chaudhary, 1977). In case of Agglomerative hierarchical clustering (AHC), standardized mean data was used for the calculation of distances between the mean of the different genotypes for characters. Euclidian distances (coefficient) were calculated and used for formation of dendrograms by unweighted pair group method using arithmetic averages (UPGMA). The number of clusters in AHC were decided by the value of Euclidian distance as dissimilarity coefficient in the dendrograms (Govindaraj et al., 2020). Therefore, the number of clusters and grouping of genotypes into different clusters varied from one method to other method of cluster analysis. However, by making consensus it is observed that the comparable type of clustering in the diversity of R line, B line and land races. Especially the small seeded land races Kuttu cumbu 1, Kuttu cumbu 2, Kuttu cumbu 3 and Pothu Cumbu were grouped in one cluster and the B line ICMB 06111 was clearly different from other clusters (Table8). Similar pattern of clustering was observed between AHC and PCA as the data used for analysis were mean data in both the cases (Chaudhary et al., 2015; Malik et al., 2017)

In the present investigation the main objective of selection of both the cluster analysis is to improve the selection criteria for the identification of diverse parents. The Mahalanobis’ D² statistics is comparatively superior as it uses the replicated data, whereas agglomerative hierarchical clustering method does not require replicated data. Principal component analysis (PCA) helps in identifying genotypes extreme for the variation which is highly useful in plant breeding (Kiprotich et al., 2015). From the principal component scores it is possible to identify the genotypes contributing highest variation for several characters and genotypes are ranked accordingly. Hence, for cluster analysis, PCA has added advantage for selection of genotypes in breeding programmes.

In the present study the main aim of employing the multivariate analysis was to identify the potential parents and classify the pearl millet genotypes comprising R lines, B lines and land races. Three types of multivariate analysis viz. D² statistics, AHC method and PCA were utilized to analyse the genetic variation present in the set of genotypes. Even though each method has its own pros and cons for synthesizing the observed data and providing classificatory analysis, all the methods were compared to identify the elite genotypes with associated quantitative traits and making a selection strategy for crop improvement. In order to apply more selection pressure and to achieve highest genetic gain from the selected potential genotypes, classificatory approaches were performed. Each method has indicated the different groupings of genotypes. However, by making consensus, the most common parents which showed diverse in all three multivariate analysis included small seeded land races like Kuttu cumbu 1, Kuttu cumbu 2, Kuttu cumbu 3 and Pothu Cumbu and also the B line ICMB 06111. Highest cluster mean for single plant yield, thousand seed weight, flag leaf length and leaf length were observed for the cluster containing the genotype Kizikuppam local constantly in both Mahalanobis’ D² statistics and AHC method of cluster analysis. PT 6706, PT 6709, Nattu Cumbu, Cumbu 2, PT 6676 and PT 6067 were the top-ranking genotypes upon PCA analysis with positive PC1 scores. High contribution of traits to total variability was by thousand grain weight, single plant yield, plant height and biological yield. Hence, there is much scope for selection of these traits among the genotypes studied for the exploitation of heterosis in hybrids and for obtaining broad spectrum of variation in segregating material for yield attributing traits.
REFERENCES

Animasaun, D. A., Morakinyo, J. A., Krishnamurthy, R. and Mustapha, O. T. 2017. Genetic divergence of Nigerian and Indian pearl millet accessions based on agronomical and morphological traits. J. Agric. Sci., 62(2): 115-131. [Cross Ref]

Athoni, B. K., Boodi, I., Pattanashetti, S. K. and Guggari, A. K. 2016. Genetic diversity for yield and its component traits in pearl millet [Pennisetum glaucum (L) R Br]. Int. J. Sci. Nat., 7(4): 795-798.

Bashir, E. M., Ali, A. M., Ali, A. M., Melchinger, A. E., Parzies, H. K. and Haussmann, B. I. 2014. Characterization of sudanese pearl millet germplasm for agro-morphological traits and grain nutritional values. Plant Genetic Resources, 12(1): 35-47. [Cross Ref]

Chaudhary, S. U., Sagar, P., Hooda, B. K. and Arya, R. K. 2015. Multivariate analysis of pearl millet data to delineate genetic variation. Forage res., 40(4): 201-208.

Dave, R. V. and Joshi, P. 1995. Divergence and heterosis for fodder attributes in pearl millet. Indian J. Genet. Plant Breed., 55: 392-397.

Drabo, I., Zangre, G. R., Sawadogo, M. and Ouedraogo, M. 2013. Genetic variability and estimates of genetic parameters in Burkina faso's pearl millet landraces. Int. J. Res. Agric. For., 3(7): 367-373.

Ghazy, M. M., Sakr, H. O. and Rajab, M. N. 2015. Estimation of genetic variability and divergence in some selected lines of pearl millet. J. Agr. Chem. &Biotechnol., 6(12): 615-626. [Cross Ref]

Govindaraj, M., Selvi, B. and Sudhir, K. I. 2011. Genetic diversity studies in indigenous pearl millet [Pennisetum glaucum (L) R Br.] accessions based on biometrical and nutritional quality traits. Indian J. Plant Genet. Resour., 24(2): 186-193.

Govindaraj, M., Yadav, O. P., Rajpurohit, B. S., Kanatti, A., Rai, K. N. and Dwivedi, S. L. 2020. Genetic variability, diversity and interrelationship for twelve grain minerals in 122 commercial pearl millet cultivars in India. Agric. Res., 9(4): 1-10. [Cross Ref]

INDIASTAT 2018. INDIASTAT database Datanet India Retrieved from https://www.indiastat.com/table/agriculture-data/2/bajraspiked-millet/17198/7267/dataaspx.

Jain, S. K. and Diwan, D. 2021. Principal component and cluster analysis for quantitative traits to identify high yielding genotypes of pearl millet [Pennisetum glaucum (I) R Br]. Forage res., 48(4): 308-314.

Kalagare, V. S., Meenakshi Ganesan, N., Iyanar, K., Chidleshware, T. and Chandrasekhar, C. N. 2021. Strategy of multiple selection indices for discrimination of potential genotypes and associated traits for yield improvement in pearl millet [Pennisetum glaucum (L.) R. Br.]. Electronic Journal of Plant Breeding. 12(3): 895-906. [Cross Ref]

Kaushik, J., Vart, D., Kumar, M., Nehra, M. and Kumar, R. 2018. Genetic diversity assessment of pearl millet maintainer lines. J. Pharmacogn. Phytochem., 7(5): 2428-2432.

Kiprotich, F., Kimurto, P., Ombui, P., Towett, B., Jeptanui, L., Henry, O. and Lagat, N. 2015. Multivariate analysis of nutritional diversity of selected macro and micronutrients in pearl millet (Pennisetum glaucum) varieties. Afr. J. Food Sci., 9(3): 103-112. [Cross Ref]

Kumar, M., Rani, K., Ajay, B. C., Patel, M. S., Mungra, K. D. and Patel, M. P. 2020. Multivariate diversity analysis for grain micronutrients concentration, yield and agro morphological traits in pearl millet (Pennisetum glaucum (L) R Br). Int. J. Curr. Microbiol. Appl. Sci., 9(3): 2209-2226. [Cross Ref]

Mahalanobis, P. C. 1936. On the generalized distance in statistics. National Institute of Science of India.

Malik, E. A., Bhardwaj, R., Goyal, M. and Kaur, J. 2017. Morpho-physiological diversity to evaluate dry summer adaptability of pearl millet. Agric. Res., 6(2): 122-129. [Cross Ref]

Murty, B. R. and Tiwari, J. L. 1967. Influence of dwarfing genes on genetic diversity in Pennisetum typhoides. Indian J. Genet. Plant Breed., 27(2): 226.

Nagendra, M.S., Selvaraju, P., Jerlin, R., Ganasesamurthy, K. and Senthil, N. 2020. Application of SSR marker diversity assessment in pearl millet inbreds. Electronic Journal of Plant Breeding, 11(04): 1181-1186. [Cross Ref]

Nehra, M., Kumar, M., Vart, D., Sharma, R. K. and Choudhary, M. 2016. DUS characterization and diversity assessment in pearl millet inbreds. Electronic Journal of Plant Breeding, 7(4): 925-933. [Cross Ref]

Patial, K. S., Gupta, S. K., Marathi, B., Danam, S., Thatikunta, R., Rathore, A., Das, R. R., Dangi, K. S. and Yadav, O. P. 2020. African and Asian origin pearl millet populations: Genetic diversity pattern and its association with yield heterosis. Crop Sci., 60: 3035-3048. [Cross Ref]

Pujar, Mahesh, Mahalingam Govindaraj, Gangaprasad,
Venkataramana. S. Kalagare et al., 2020. Genetic variation and diversity for grain iron, zinc, protein and agronomic traits in advanced breeding lines of pearl millet \([Pennisetum glaucum (L) R. Br.]\) for biofortification breeding. \textit{Genet. Resour. Crop Evol.} \textbf{67}: 2009-2022. [Cross Ref]

Ramya, A. R., Ahamed, M. L. and Srivastava, R. K. 2017. Genetic diversity analysis among inbred lines of pearl millet \([Pennisetum glaucum (L) R Br]\) based on grain yield and yield component characters. \textit{Int. J. Curr. Microbiol. Appl. Sci.}, \textbf{6(6)}: 2240-2250. [Cross Ref]

Reddy, K. N., Rao, S. A. and Mengesha, M. H. 1996. Diversity in pearl millet germplasm from Central African Republic. \textit{Genet. Resour. Crop Evol.}, \textbf{43(4)}: 303-308. [Cross Ref]

Sankar, S. M., Satyavathi, C. T., Singh, S. P., Singh, M. P., Bharadwaj, C. and Barthakur, S. 2014. Genetic diversity analysis for high temperature stress tolerance in pearl millet \([Pennisetum glaucum (L) R. Br.]\). \textit{Indian J. Plant Physiol.}, \textbf{19(4)}: 324-329. [Cross Ref]

Santos, R., Neves, A. L., Pereira, L. G., Verneque, R., Costa, C. T., Tabosa, J., Scherer, C. and Gonçalves, L. 2017. Divergence in agronomic traits and performance of pearl millet cultivars in Brazilian semiarid region. \textit{Grassl. Sci.} \textbf{63(2)}: 118-127. [Cross Ref]

Shanmuganathan, M., Gopalan, A. and Mohanraj, K. 2006. Genetic variability and multivariate analysis in pearl millet \([Pennisetum glaucum (L) R. Br.]\) germplasm for dual purpose. \textit{J. Agric. Sci.}, \textbf{2}: 73-80. [Cross Ref]

Sharma, B., Chugh, L., Singh, V. K., Shekhar, C. and Tanwar, N. 2020. Characterization of rancidity indicators in selected pearl millet genotypes by multivariate analysis. \textit{Plant Arch.}, \textbf{20}: 229-235.

Singh, R. K. and Chaudhary, B. D. 1977. Biometrical methods in quantitative genetic analysis. Kalyani publishers. 229-252.

Singh, S. and Gupta, P. K. 1979. Genetic divergence in pearl millet. \textit{Indian J. Hum. Genet.}, \textbf{32}: 210-215.

Singh, Y. P., Kumar, A. and Chauhan, B. P. S. 1981. Genetic divergence in pearl millet. \textit{Indian J. Genet. Plant Breed.}, \textbf{41(2)}: 186-190.

Statistical Handbook of Tamil Nadu 2019. Directorate of economics and Statistics, Government of Tamil Nadu (GOI).

Sumathi, P., Lalithkannan, R. and Revathi, S. 2016. Genetic analysis and diversity studies in pearl millet \([Pennisetum glaucum (L) R Br]\). \textit{Electronic Journal of Plant Breeding}, \textbf{7(4)}: 1014-1019. [Cross Ref]

Swamynathan, S., Kumar, M. H., Reddy, D. M. and Latha, P. 2020. Genetic divergence studies in pearl millet \([Pennisetum glaucum (L) R. Br.]\). \textit{Electronic Journal of Plant Breeding}, \textbf{11(01)}: 76-80. [Cross Ref]

Upadhyay, M. K. and Murty, B. R. 1970. Genetic divergence in relation to geographical distribution in pearl millet. \textit{Indian J. Genet. Plant Breed.} \textbf{30}: 704-715.

Yadav, O. P., Rai, K. N., Yadav, H. P., Rajpurohit, P. S., Gupta, S. K., Rathore, A. and Karjagi, C. G. 2016. Assessment of diversity in commercial hybrids of pearl millet in India. \textit{Indian J. Plant Genet. Resour.}, \textbf{29(2)}: 130-136. [Cross Ref]

Zhang, X., Gu, H., Ding, C., Zhong, Xiaoxian, Zhang, Jianli and Xu, N. 2010. Path coefficient and cluster analyses of yield and morphological traits in \textit{Pennisetum purpureum}. \textit{Trop. Grassl.} \textbf{44}: 95-102.