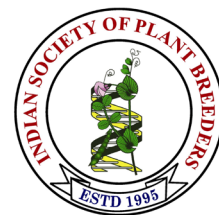# Electronic Journal of Plant Breeding

Research Article

Principal components of genetic diversity and association studies for yield related traits in pearl millet [*Pennisetum glaucum* (L.) R. Br.]

R. Narasimhulu[1*], C. Tara Satyavathi[2], B. Sahadeva Reddy[1] and B. C. Ajay[3]

[1]Agricultural Research Station (ANGRAU), Ananthapuramu, Andhra Pradesh
[2]All India Coordinated Research Project on Pearl Millet, Mandor, Rajasthan
[3]Regional Research Station, ICAR-Directorate of Groundnut Research, Ananthapuramu, Andhra Pradesh
*E-Mail: narsi46@gmail.com

Abstract
One sixty-eight pearl millet germplasm accessions were evaluated using multivariate techniques, principal component and cluster analysis to explore the extent of genetic diversity and association among the yield and yield contributing traits. The first three principal components PC1, PC2 and PC3 with eigen values more than one explains 25.27, 22.15 and 13.38 per cent, respectively, with a cumulative effect of 60.80 per cent of the total variation. Based on factor loading of three PC's and association analyses explained that the traits *viz*., number of productive tillers per plant, 1000-seed weight, panicle diameter and plant height function as excellent selection indicators that exhibit positive loading toward genetic diversity. In cluster analysis all the germplasm accessions under the study were sorted into seven distinct clusters. Cluster I having the highest number of lines followed by cluster IV due to similar ancestry of parental lines. The germplasm accessions 4129, 4163, 4157, 5007, 5058, 4104, 4105, 4146, 4150, 4140, 4123 and ICMV 221 exhibit diversity in both cluster and PCA analyses, indicating that utilizing these lines for hybridization procedure to harness heterosis and to select superior segregants in pearl millet.

Keywords: Pearl Millet, PCA, Biplot, Trait Association and Grain Yield

INTRODUCTION

Pearl millet (*Pennisetum glaucum* (L.) R. Br.) is a highly cross-pollinated, diploid (2n=14) annual C_4 crop native to central tropical Africa. It is widespread across Africa and India's arid and semi-arid areas, where it was traditionally cultivated. Pearl millet exhibits a tremendous amount of diversity at both phenotypic and genotypic levels and it has protogyny and anemophily mechanisms, which meet the biological criteria for hybridization. In the process of genetic improvement of the crop, genetic diversity among germplasm based on the contribution of various qualitative and quantitative traits plays a key role, as it allows for the identification of the most divergent parents for further

utilization breeding programme. Therefore, there is a need to evaluate available germplasm for their genetic diversity. Apart from genetic divergence, correlation coefficient indicates which traits have a direct impact on seed yield and provides background for selection and development of new genetically diverse inbred.

Multivariate statistical techniques like cluster analysis (CA) and principal component analysis (PCA) based on morphological or molecular markers assists the breeders in selecting appropriate genotypes that fulfil the breeding programme objectives (Mohammadi and

Prasanna, 2003). PCA has an edge over the other as it eliminates multicollinearity among the independent variables and shows the relevance of the largest contributor to the total variance at each differentiation axis. Previously, multivariate analysis was extensively employed for genetic diversity analysis in pearl millet (Jyoti *et al.,* 2018; Singh *et al.,* 2018; Kaushik *et al.,* 2018a; Rasitha *et al.,* 2020; Sanjana Reddy *et al.,* 2021). The goal of this study was to use PCA to assess the relative contribution of various traits to total variability in a collection of pearl millet germplasm accessions, as well as cluster analysis to group these accessions into distinct clusters and establish an acceptable selection criterion for selecting the best parental lines for a trait-specific line development programme

MATERIALS AND METHODS
The experiment was carried out at the Agricultural Research Station, ANGRAU, Ananthapuramu (14° 41' N, longitude: 77 ° 40' E and 373 m above mean sea level) in Andhra Pradesh, India, in the scarce rainfall zone. This district has an annual rainfall average of about 553 mm and is characterized by low and erratic rainfall in terms of area, time and distribution during the season. During *Kharif* 2020, 168 genotype, comprising 163 germplasm accessions and five popular varieties serving as checks were planted in an Augmented design. Each accession representing one row of 4 m length spaced 50 cm apart with 15 cm between hills. The field was uniformly fertilized with a basal dose of 30 kg Nitrogen, 20 kg P_2O_5 and 20 kg K_2O per hectare just before seeding and a dose of 30 kg Nitrogen per hectare 35 days after seeding. Standard cultural and agronomic practices were adopted to achieve

good crop growth. Five competitive plants per genotype were selected randomly for recording observations on different characters *viz*., plant height (cm), number of productive tillers per plant, panicle length (cm) and panicle diameter (cm), while observations on days to 50% flowering, days to maturity and grain yield (kg/ha) were recorded on plot basis. 1000-Seed weight, a sample of 1000 grains were counted randomly from the threshed seed and the weight is recorded in grams. R-software was used to compute Pearson's correlation coefficients for the nine individual traits (R Development Core team 2020). Variability among the genotypes was calculated using principal component analysis (PCA) following Davis (1986) using a correlation matrix in the R program (R core team 2018). The PCs with eigen value >1.0 were considered determining the agro-morphological variability in the genotypes (Kaiser, 1960). For PCA analysis data was normalized using square root transformation and PC-1 and PC-2 scores were used to create a scatter plot of traits.
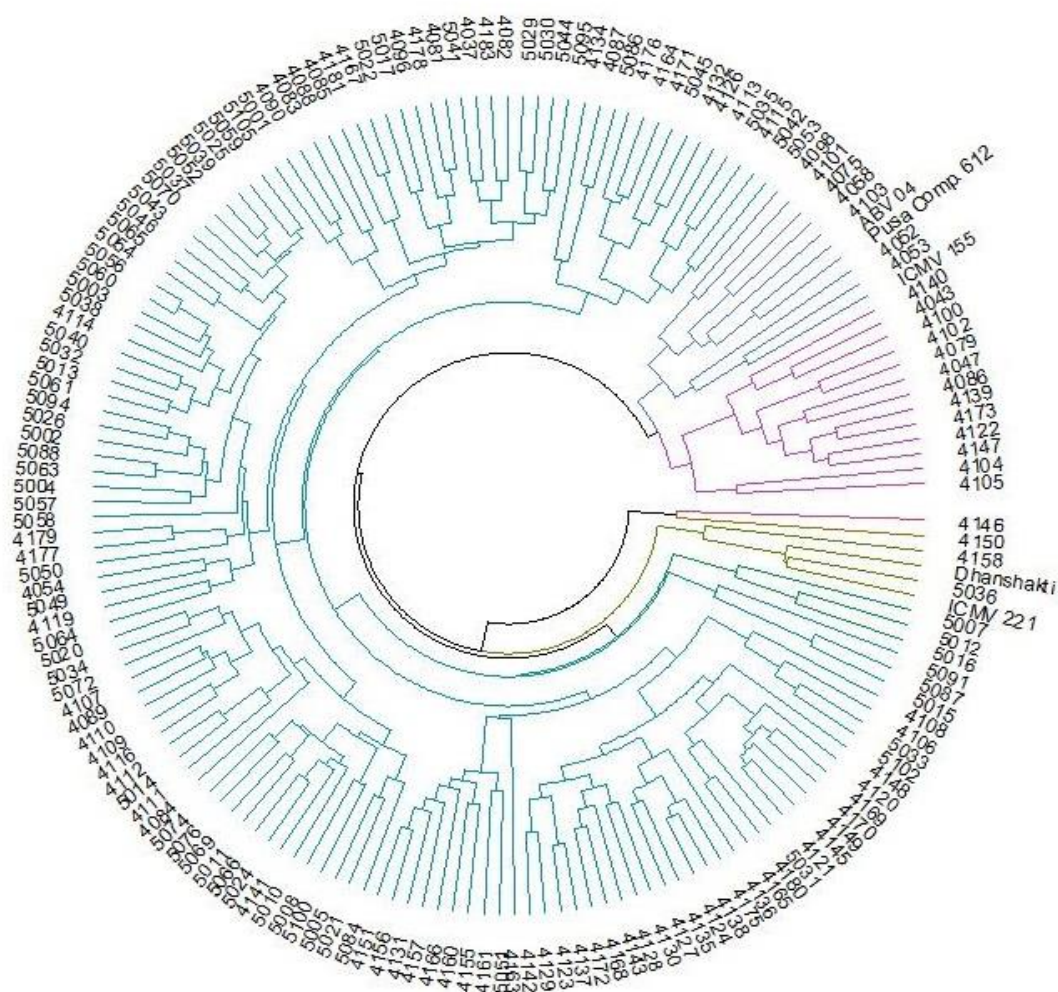
RESULTS AND DISCUSSION
The degree of divergence between the selected parents determines the effectiveness of producing highly heterotic hybrids and provides ample scope for selection of beneficial recombinants among the segregants. Cluster analysis is a method to classify the total entities into a number of different groups based on a set of deliberately chosen criteria, while keeping related accessions together in the same group. Based on different traits, all 163 germplasm accessions and five checks were classified into seven main groups (**Table 1**) using cluster analysis. These seven primary clusters further divided into sub clusters **Fig. 1**.

Table 1. Grouping of pearl millet germplasm accessions into different clusters

S. No	Cluster number	Number of genotypes	Genotypes
1	I	139	5012, 5016, **5007**, 4096, 4183, 5014, 5034, 4082, 4181, 4155, 5020, 4089, 5029, 5010, 4084, 4112, 4085, 5044, 4167, 4178, 4111, 5042, 5057, 4107, 4081, 4141, 4116, 5041, 4160, 5072, 5049, 5004, 5030, 5021, 5035, 4054, 4166, 5005, 5024, 4156, 4110, **5058**, 5022, 5053, 4164, 5032, 4169, 5063, 5003, 4037, 5064, 5040, 5091, 4115, 5011, 5017, 5070, 4171, 5066, 4170, 5051, 5026, 5052, 5076, 5054, 5069, 5095, 4083, 4109, 4130, 4088, 5074, 5025, 4148, 4145, 4180, 5059, 4165, 4151, 4113, 4135, 4131, 4168, 4120, 5037, 5039, 5094, 5100, 4121, 4176, 4108, 5065, 4175, **4163**, 4149, 4132, 4172, 4126, 4087, 4177, 5088, 5086, 4127, 5008, 5060, 5038, 5061, 5056, 4106, 5013, 4179, 4114, 4138, 5046, 5001, 5043, **4157**, 5084, **4129**, 4124, 4090, 5087, 4128, 4143, 4119, 5002, 4136, 4161, 5105, 4142, 5045, 4137, 5050, 5031, **4123**, 5015, 4134, 5033, 5102
2	II	7	4098, 4101, 4075, 4058, 4103, ABV 04, Pusa Comp. 612
3	III	3	4052, 4053, ICMV 155
4	IV	11	**4140**, 4043, 4100, 4102, 4079, 4047,4086, 4139, 4173, 4122, 4147
5	V	2	**4104, 4105**
6	VI	2	4146, 4150
7	VII	4	4158, Dhanshakti, 5036, **ICMV 221**

**Note: Bolded accessions showed diverse nature in PCA scatter plot also**

Fig. 1. Dendrogram depicting clustering pattern of 168 germplasm accessions obtained by cluster analysis in Pearl millet

Out of seven major clusters, cluster I was the largest, with 139 accessions, followed by cluster IV with eleven accessions, cluster II with five accessions along with two recently released varieties ABV 04 and Pusa Comp. 612; cluster VII with two accessions along with two popular varieties Dhanshakti and ICMV 221 and cluster III with two accessions with one popular variety ICMV 155. Clusters V and VI, on the other hand, each included two accessions, suggesting a significant degree of variability among the accessions. In pearl millet, Ramya *et al.* (2017), Singh *et al.* (2017), Kaushik *et al.* (2018b), and Kumar *et al.* (2020) previously described clustering of germplasm accessions based on quantitative data to identify prospective parents for heterotic expression of yield components. Selection of parents from these cluster I, V and VI will serve as a potential breeding material for improving the yield attributing traits in hybridization programme. Similar results were reported by Singh *et al.* (2017) and Kumari *et al.* (2016). Early

flowering genotypes mostly suitable for cultivation in drought prone area Sankar *et al.* (2014). Considering this point, individuals belonging to cluster I used for developing early flowering lines with higher yield.

Yield amelioration is the ultimate aim of any plant breeder in crop improvement programme which can be achieved through gaining insight knowledge on association of component traits with grain yield. The correlation matrix is a technique for concurrently analysing the interdependence of several variables simultaneously and these associated traits enables the selection of superior genotypes *via* phenotypic selection of easily heritable traits. In correlation matrix (**Fig. 2**) on the lower half, the vicariate schematic model plots with a line of best fit can be seen and upper half represents, the values of the correlation coefficients indicated with significance level as stars. Each traits distribution is represented on the diagonal whether it is distributed normally or not.
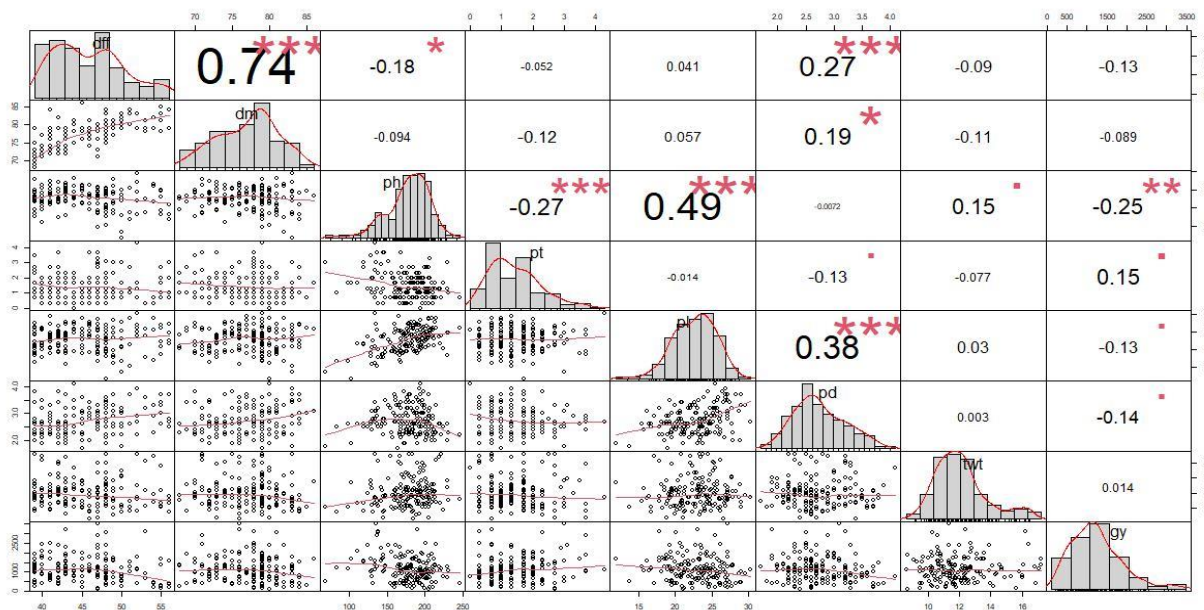
Fig. 2. Correlation for yield and yield contributing traits in Pearl millet
dff: Days to 50 % flowering, dm: Days to maturity, ph: Plant height , pt: Productive tillers , pl: Panicle length, pd: Panicle diameter, twt: 1000 - grain weight, gy: Grain yield

In the present study, the number of productive tillers per plant had a significant positive association with grain yield (0.15*), while the 1000-grain weight (0.014) had a non-significant positive relationship. Previously, Annamalai *et al.* (2020) and Narasimhulu *et al.* (2021a) found a favourable connection between grain yield and productive tillers and 1000-grain weight. Plant height (-0.25**), panicle length (-0.13*) and panicle diameter (-0.141*), on the other hand, were shown to have a significant and negative relationship with grain yield, implying that greater scope to select high yielders with short stature, which may be useful for developing the cultivars for lodging tolerance and suitable for mechanical harvest. Days to 50% flowering and days to maturity had a non-significant negative association with grain yield, which is desirable for development of early maturity hybrids with high grain yield.

From the inter-association studies, it was observed that days to 50% flowering exhibited significant and positive relationship with days to maturity (0.74**) which was also confirmed by Narasimhulu *et al.* (2021b). Early flowering/ maturation accessions exhibited fewer productive tillers per plant (-0.12) and had larger panicle diameters (0.19*). Rasitha *et al.* (2019) and Narasimhulu *et al.* (2021a) found similar findings for the number of productive tillers per plant and panicle diameters. The length of a panicle, on the other hand, is proportionate to the plant's height (0.49**) and panicle diameter (0.38**). These findings agree with the results of the Narasimhulu *et al.* (2021b). In a nutshell,

the majority of the germplasm accessions were found to have small stature with a higher number of productive tillers in the early maturing type background, suggesting that there is greater potential for developing trait specific parental lines for future breeding programmes.

Principal Component Analysis was used as a reductionist approach of the multivariate data, to assess the relevance and contribution of each component to total variance. The significance of traits towards the principal components can be realized from the respective eigen values presented in **Table 2**. In practise, the PCA approach is used for genetic diversity and genotype grouping using biplot diagrams, with applications in crop breeding for selection of superior genotypes (Singh *et al.*, 2018; Kaushik *et al.*, 2018a; Rasitha *et al.*, 2020). The results of PCA analysis indicated that the first three components had eigen values more than one and contributed around 60.80 per cent of total accumulated variability for the eight different morphological traits investigated among the 168 germplasm accessions.

The first principal component (PC1) accounted for 25.27 per cent of the total variation, largely due to grain yield (0.270), number of productive tillers (0.209) and 1000-grain weight (0.068) indicating that these traits were important contributors to variation among the accessions and high yielding lines could be distinguished based on these traits (**Table 3**). The existence of both positive and negative loading in a single component indicates the presence

Table 2. Principal components indicating eigen values, proportion of variation and total variation across axis

Principal component axis	1	2	3	4	5	6	7	8
Eigenvalues	1.42	1.33	1.03	0.98	0.91	0.89	0.55	0.49
Proportion of Variance (%)	25.27	22.15	13.38	12.12	10.27	10.01	3.78	3.03
Cumulative variation (%)	25.27	47.43	60.80	72.92	83.19	93.19	96.97	100.00

Table 3. Principal component analysis for quantitative traits in pearl millet accessions – non rotated loadings

| Particulars	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
Days to 50 % flowering	-0.553	0.338	-0.073	0.092	0.161	-0.150	-0.048	-0.718
Days to maturity	-0.545	0.303	-0.142	0.056	0.101	-0.340	-0.025	0.680
Plant height	-0.103	-0.619	-0.084	-0.124	0.080	-0.441	-0.615	-0.069
Productive tillers	0.209	0.249	0.656	0.034	0.620	-0.053	-0.270	0.062
Panicle length	-0.293	-0.464	0.509	0.076	-0.010	-0.246	0.610	-0.026
Panicle diameter	-0.424	-0.114	0.362	0.245	-0.336	0.586	-0.388	0.096
1000 - grain weight	0.068	-0.237	-0.312	0.816	0.394	0.120	0.068	0.031
Grain yield	0.270	0.252	0.224	0.490	-0.552	-0.496	-0.137	-0.061

of positive and negative correlation trends between the components and the variables. Hence, these characters which load high positively or negatively contributed more to the diversity (Bhanupriya *et al.,* 2014). Five traits *viz.,* 50% flowering, days to maturity, plant height, panicle length and panicle diameter all contributed negatively to the first component and showed a negative association, implying that high yielding accessions with larger seed size and more number of productive tillers per plant had shorter plant stature and panicle length in the background of early maturity, this showed that these accessions can escape from the terminal drought or water shortage and may useful for developing the genotypes suitable for mechanical harvest as conformed the same with the association studies.

Second principal component (PC2) accounted for 22.15 per cent of the total variation contributed through days to 50% flowering (0.338), days to maturity (0.303), grain yield (0.252) and number of productive tillers (0.249). Grain yield was shown to be increased by increasing the number of productive tillers per plant with a late maturity background. The third principal component (PC3) contributed 13.38 per cent through number of productive tillers (0.656), panicle length (0.509), panicle diameter (0.362) and grain yield (0.224). According to PC3, grain yield was enhanced by a greater number of productive tillers per plant and increasing the length and girth of the panicles. Similarly, PC4, PC5, PC6, PC7 and PC8 each provided 12.12, 10.27, 10.01, 3.78 and 3.03 per cent of the total variation in the germplasm, respectively. The prominent characters identified in a particular principal component as prime contributors to total variability have the tendency to hang together and can be used successfully for selection in crop improvement programmes. In all the eight components traits like number of productive

tillers per plant, 1000-seed weight, panicle diameter and in few cases traits like plant height and panicle length contributed positively to the total variation. Hence these traits could be exploited maximum in order to get superior hybrids with higher yield in pearl millet.

A PC biplot **Fig. 3** showed that variables and genotypes are superimposed on the plot as vectors. The contribution of these variables in the variation of the genotypes is represented by the distance of each variable with respect to PC1 and PC2. The biplot diagram shows that genotypes are equally scattered around the vectors, resulting in different distinct groups. The accessions 4129, 4163, 4157, 5007, 5058, 4104, 4105, 4140, 4123 and ICMV 221 were placed in different quadrants at extreme ends of the plot. This suggests that these genotypes highly varied from the rest and they are diverse from each other for all the characters. Hence, these accessions were identified as most divergent and distinct and could be used as candidate parents in crop improvement programme to develop promising hybrids. Another consideration of the biplot diagram is the angles of vectors. Kroonenberg (1995) concluded that the angle of vectors shows correlations among the traits. The biplot showed that grain yield and no. of productive tillers; day to 50 % flowering and day to maturity vectors were in overlapping to each other with a small angle, indicates that they had positive associations. Day to 50 % flowering, day to maturity, grain yield, panicle diameter, panicle length and plant height contributed maximum towards variability in germplasm.

The findings of this study revealed that there is a lot of genetic diversity among the germplasm accession and the possibility of developing potential early maturing high yielder with larger seed size and more productive tillers per plant with shorter plant stature and larger panicle.
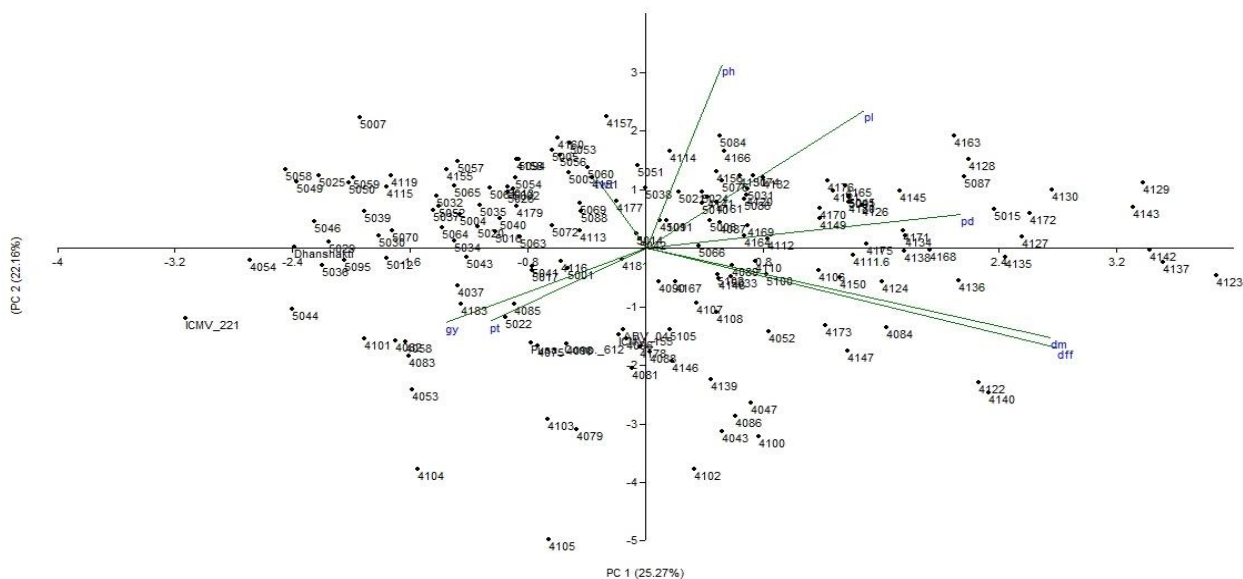
Fig. 3. Biplot indicating the best genotype for the target traits

This plant type is critical in the current changing climatic situation for developing robust lines with the capacity to sustain terminal drought or water scarcity, as well as being suited for mechanical harvest. Based on both cluster and PCA analyses, the germplasm accessions 4129, 4163, 4157, 5007, 5058, 4104, 4105, 4146, 4150, 4140, 4123, and ICMV 221 were identified as prospective parents for heterotic expression and to generate potential breeding material for improving trait specific materials to enhance productivity in the pearl millet.

REFERENCES

Annamalai, R., Aananthi, N., Arumugam Pillai, M. and Leninraja, D. 2020. Assessment of variability and character association in pearl millet [*Pennisetum glaucum* (L.) R.Br.]. *Int. J. Curr. Microbiol. App. Sci.,* **9**(06): 3247-3259. [Cross Ref]

Bhanupriya, B. Das, Satyanarayana, N.H., Mukherjee, S. and Sarkar, K. K. 2014. Genetic diversity of wheat genotypes based on principal component analysis in Gangetic alluvial soil of West Bengal. *J. Crop and Weed*, **10**(2):104-107.

Davis, J. C. 1986. Statistics and data analysis in geology. John Wiley and Sons

Jyoti, K., Dev Vart, Mukesh Kumar and Anil Kumar. 2018. Genetic diversity analysis among maintainer lines of pearl millet [*Pennisetum glaucum* (L.) R. Br.] based on grain yield and yield component characters. *Chem Sci Rev Lett.,* **7**(28): 978-981.

Kaiser, H. F.1960. The application of electronic computers to factor analysis. Educ. Psychol. Meas. **20:** 141–151. [Cross Ref]

Kaushik, J., Vart, D., Kumar, M. and Kumar, A. 2018a. Genetic diversity analysis among maintainer lines of pearl millet [*Pennisetum glaucum* (L.) R. Br.] based on grain yield and yield component characters. *Chem Sci Rev Lett.,* **7**(28): 978- 981.

Kaushik, J., Vart, D., Kumar, M., Kumar, A. and Kumar, R. 2018b. Phenotypic diversity in Pearl Millet [*Pennisetum glaucum* (L.) R. Br.] germplasm lines. *Int. J. Chem. Stud.,* **6**(5): 1169- 1173.

Kroonenberg, P.M. 1995. Introduction to biplots for G × E tables. Department of mathematics, Res. report 51. University of Queensland, Australia.

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. App. Sci.,* **9**(3): 2209-2226. [Cross Ref]

Kumari, J., Bag, M., Pandey, S., Jha, S., Chauhan, S., Jha, G. K., Gautam, N. and Dutta, M. 2016. Assessment of phenotypic diversity in pearl millet [*Pennisetum glaucum* (L.) R. Br.] germplasm of Indian origin and identification of trait-specific germplasm. *J. Crop Past. Sci.,* **67**(12): 1223-1234. [Cross Ref]

Mohammadi, S.A. and Prasanna, B.M. 2003. Analysis of genetic diversity in crop plants, salient statistical tools and considerations. *Crop Sci.,* **43**: 235-48. [Cross Ref]

Narasimhulu, R., Sahadeva Reddy, B., Tara Satyavathi, C. and Ajay, B.C. 2021a. Performance, genetic variability and association analysis of pearl millet

yield attributing traits in Andhra Pradesh's arid region. *Chem. Sci. Rev. Lett.,* **10**(38): 177-182.

Narasimhulu, R., Sahadeva Reddy, B., Tara Satyavathi, C. and Ajay, B.C. 2021b. Analysis of the performance, variability and nature of Pearl Millet productive traits in the rainfed environment of Andhra Pradesh. *Electronic Journal of Plant Breeding,* **12**(3): 766-771. [Cross Ref]

R Development Core Team. 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.

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. *Int. J. Curr. Microbiol. App. Sci.,* **6**(6): 2240-2250. [Cross Ref]

Rasitha, R., Iyanar, K., Ravikesavan, R. and Senthil, N. 2020. Assessment of genetic diversity in parental lines of pearl millet [*Pennisetum glaucum* (L). R. Br] for yield and yield related traits. *Int. J. Curr. Microbiol. App. Sci.,* **9**(12): 1575-1582. [Cross Ref]

Rasitha, R., Iyanar, K., Ravikesavan, R. and Senthil, N. 2019. Studies on genetic parameters, correlation and path analysis for yield attributes in maintainer and restorer lines of pearl millet [*Pennisetum glaucum* (L.) R. Br.]. *Electronic Journal of Plant Breeding,* **10**(2): 382-388. [Cross Ref]

Sanjana Reddy, P., Satyavathi, C.T., Khandelwal, V., Patil, H.T., Gupta, P.C., Sharma, L.D., Mungra, K.D., Singh, S.P., Narasimhulu, R., Bhadarge, H.H., Iyanar, K., Tripathi, M.K., Yadav, D., Bhardwaj, R., Talwar, A.M., Tiwari, V.K., Kachole, U.G., Sravanti, K., Shanthi Priya, M., Athoni, B.K., Anuradha, N., Govindaraj, M., Nepolean, T. and Tonapi, V.A. 2021. Performance and stability of pearl millet varieties for grain yield and micronutrients in arid and semi-arid regions of India. *Front. Plt Sci.*, **12**: 670201. [Cross Ref]

Sankar, S.M., Satyavathi, C.T., Singh, S., 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]. *Ind. J. Plt. Phys.*, **19**(4):324- 329. [Cross Ref]

Singh, O.V., Gowthami, R., Singh, K. and Neelam S. 2018. Genetic divergence studies in pearl millet germplasm based on principal component analysis. *Int. J. Curr. Microbiol. App. Sci.,* **7**(06): 522-527. [Cross Ref]

Singh, O M., Gowthami, R., Neelam S. and Singh, K. 2017 Genetic divergence studies in pearl millet [*Pennisetum glaucum* (L.).R.Br] germplasm using Mehalanobis D^2 analysis over five years in hot arid climate of Rajastan. *Int. J. Curr. Microbiol. App. Sci.,* **7**:2319-7706. [Cross Ref]