Plant Archives Vol. 26, Supplement 1, 2026 pp. 2247-2251 e-ISSN:2581-6063 (online), ISSN:0972-5210

Plant Archives

Journal homepage: http://www.plantarchives.org
DOI Url : https://doi.org/10.51470/PLANTARCHIVES.2026.v26.supplement-1.303

ASSESSMENT OF GENETIC DIVERSITY USING CLUSTER AND PRINCIPAL COMPONENT ANALYSIS FOR YIELD AND ITS CONTRIBUTING TRAITS OF BREAD WHEAT (*TRITICUM AESTIVUM* L.)

V. K. Jaiswal[1], Prem Kumar[1], Vijay Kumar Yadav[1]*, Vaibhav Chauhan[1], Krishan Kumar[2], Suraj Muttepawar[1], Amit Kumar[1], Shivanshu Shekhar[1] and Sachin Pal[1]
[1]Department of Genetics and Plant Breeding, Chandra Shekhar Azad University of Agriculture & Technology, Kanpur (U.P.) India-208002
[2]Department of Vegetable science, Chandra Shekhar Azad University of Agriculture & Technology, Kanpur (U.P.) India-208002
*Corresponding authors E-mail: vkyadu@gmail.com

(Date of Receiving : 20-09-2025; Date of Acceptance : 02-12-2025)

ABSTRACT

An experiment was carried out to assess genetic diversity by cluster and principal component analysis (PCA) for yield and its contributing characters in 40 bread wheat genotypes at Crop Research Farm, Nawabganj, Chandra Shekhar Azad University of Agriculture & Technology, and Kanpur (U.P.) during *Rabi* season 2020-21. Principal component analysis (PCA) indicated that four components (PC1 to PC4) accounted for 77.5 % of the total variation among traits in bread wheat cultivars. Out of total principal components retained PC1, PC2 and PC3 with values of 25.70%, 23.55% and 19.30%, respectively contributed most of the total variation. The first principal component had high positive loading for 3 characters out of 10 viz. days to 50 % flowering, days to maturity and weight of 1000 grains which contributed more to the diversity. Fruitful information extracted from K mean clustering that genotypes of cluster IV and VI may be crossed with any superior genotype of cluster VII and V may be proved better in hybridization program for higher yield. The result of present study could be exploited in planning and execution of future breeding programme in wheat for generating heterotic hybrids.

***Keywords* :** Wheat, PCA, Genetic diversity, Cluster analysis, grain yield.

Introduction

Wheat (*Triticum aestivum* L., 2n=42) is one of the most important cereals in the world and was one of the first crops to be domesticated approximately 10,000 years ago (Harlan and Zohary, 1966). It is a staple food for around two billion people (36% of the global population) worldwide. Wheat yield, stability, and disease resistance have continuously improved as breeding progresses (Kiszonas *et al.*, 2018). Breeding projects have also effectively addressed global food security concerns, throughout the process in a doubling of wheat production in 20 years (Curtis and Halford, 2014). However, the continued effects of global climate change need the creation of crop types capable of adapting to changing environmental factors (Lopes *et al.*, 2015). In order to improve wheat, we must plug into the genetic diversity in the wheat gene pool to identify different genotypes for use in wheat breeding. Estimating genetic distance is one of the acceptable strategies for parental selection in wheat hybridization programs. Appropriate parent selection is important in crossing nurseries to maximize genetic recombination and potentially increase yield (Islam *et al.*, 2004).Cluster analysis, PCA, and component analysis are some suitable techniques for studies of genetic diversity, parental selection, and the relationship between environment and genotype (Eivazi *et al.*, 2008). Principal component analysis (PCA) and cluster analysis are two important methods for analysing genetic variation that differ from one another. The advantage of PCA over cluster analysis is that each germplasm line may be given to a single group (Mohammadi, 2002), while cluster analysis is a suitable technique for evaluating ancestral connections (Mellingers, 1972). The aim of the present research is to use the statistical methods of principal component analysis and cluster analysis to determine the productive genetic diversity among wheat genotypes cultivated in Uttar Pradesh's central plain zone.

Materials and Methods

K-cluster mean analysis and principal component analysis (PCA) for yield and its ten contributing characters were used in the experiment to evaluate genetic diversity in 40 bread wheat genotypes at Crop Research Farm, Nawabganj, Chandra Shekhar Azad University of Agriculture & Technology, and Kanpur (U.P.) during the 2020–21 Rabi season. Each genotype was sown in two rows of 5.0 m length with a spacing of 22.5×5 cm, and the experiment was conducted using a Randomized Complete Block Design (RCBD) with three replications. To cultivate a healthy crop, recommended agro-techniques and safety precautions were observed. Data were recorded on 10 characters viz., Days to 50% flowering, Days to maturity, Number of productive tillers per plants, plant height, Ear length, Number of spikelets per spike, Number of grains per spikes, 1000-grain weight (gm), Chlorophyll content and Grain yield per plant, Data of randomly selected five plants of each germplasm line were averaged replication wise and mean data was used for statistical analysis. The statistical tool SAS (Statistical Analyses Software) was utilized for principal component and K cluster mean analysis Hartigan & Wong (1979).

Result and Discussion

The principal component analysis, a type of multivariate analysis used to study of wheat genotypes, reflects the importance of the largest contributor to total variance along each axis of differentiation. The eigen values are frequently used to determine how many factors to retain. The sum of the eigen values is often equal to the number of variables (Sharma, 1998). Three major components had more than one eigen value. The PCA scores for 40 genotypes were computed and presented graphically in Fig 1. Variations were found for ten distinct traits in 40 core elite wheat germplasm. The primary, secondary, and tertiary axes of differentiation accounted for the majority of the variance captured by the first three canonical vectors. In the principal component analysis, three of the four principal components had more than one eigen value and demonstrated 68.56 percent variability. As a result, the three axes explained the majority of the variation in germplasm (Table 1). The first main component vector explained 25.70% of the variation. The second main component vector explained approximately 23.55% of the variation. The third major component contributed approximately 19.30% of the total variation. It was concluded that the first PCA included greatest variation. So, selecting genotypes from this PCA will be beneficial. In the current study, genotype separation into separate clusters was caused by the comparatively high contribution of a few characters rather than the minor contribution of each character. The positive and negative loadings indicate the presence of positive and negative correlation trends between the components and variables. As a result, the characteristics with a high load, either positively or negatively, contributed more to variety or were the most distinctive characteristics among the clusters. Accordingly, Table 1 shows that the first principal component (PCA-I) had a high positive component loading from weight of 1000 grains and Days to maturity, while number of grains per spike, grain yield per plant, number of spikelets per spike, ear length per plant, and plant height had a high negative loading. Days to 50% flowering, number of spikelets per spike, ear length per plant, and number of grains per spike showed high positive component loading in the second principal component (PCA-II). This was followed by high negative loading from plant height, weight of 1000 grains, and productive tiller per plant. The number of spikelets per spike, chlorophyll content, and days to 50% flowering were the main contributing factors for the diversity in the third principal component (PCA-III). Days to maturity and ear length per plant displayed strong negative values, whereas days to 50% flowering and days to maturity showed large positive loadings. Among the ten variables, it was evident from the three PCAs that the weight of 1000 grains, days to maturity, days to 50% flowering, number of spikelets per spike, ear length per plant, and chlorophyll content all showed positive loadings. The weight of 1000 grains and days to maturity, which showed large positive loadings in the first PC, clearly had the most impact on the genotypic clustering pattern of all the features in the experiment result is in conformity with the findings of Hailegiorgis *et al.* (2011), Mishra *et al.* (2015) and Khan *et al.* (2015).

The K Cluster mean analysis revealed that Cluster V had 11 genotypes, Cluster III had 9 genotypes, Cluster II and VI had 7 genotypes each, Cluster VII had 3 genotypes, Cluster IV had 2 genotypes, and Cluster I had one genotype. K-cluster mean of principal component analysis for all ten characters listed in Table 3. The individual cluster showed supremacy for several attributes as shown by the cluster mean value in Table 3. Cluster IV showed the lowest K-cluster means for days to 50% flowering, ear length per plant, number of spikelets per plant, number of grains per spike, chlorophyll content, and grain yield per plant. Cluster I has the lowest K-Cluster mean in terms of days to maturity, while Cluster II had the lowest K-Cluster mean in terms of plant height and productive tillers per plant. Cluster VII showed the lowest K-Cluster mean for 1000 grain weight. The cluster mean value in Table 3 indicates how each unique cluster represented supremacy for different attributes. Cluster I has the highest K-Cluster means for plant height, ear length per plant, number of grains per spike, and chlorophyll content. The K-Cluster Means for days to maturity was greatest for Cluster II. Cluster IV showed the greatest K-Cluster mean at 1000 grain weight. Cluster V has the greatest K-cluster mean for productive tillers per plant. However, for improving the grain yield, the diverse clusters, i.e. cluster IV and cluster VI may be used in wheat hybrid breeding programs. Result is in conformity with the findings of Maqbool*et al.*,(2010).

Conclusion

The principal component analysis identified key variables in wheat accessions affecting agronomic traits, including days to flowering, days to maturity, productive tillers per plant, plant height, Ear length per plant, number of spikelets per spike, number of grains per spike, 1000 grain weight, Chlorophyll content and grain yield per plant. These variables can be used for effective selection. For improving grain yield per plant, Diverse parents for improvement in yield per plant based on genetic distance, K-cluster mean and mean performance was found to be EC 577050 and IC 443766.Fruitful information extracted from K mean clustering that genotypes of cluster IV and VI may be crossed with any superior genotype of cluster VII and V to may be proved better in hybridization program for higher yield.

Table 1: Canonical Root Analysis (P.C.A)

		1 Vector	2 Vector	3 Vector	4Vector
	Eigen Value (Root)	2.57018	2.35550	1.93098	0.89539
	%Var. Exp.	25.70176	23.55501	19.30984	8.95393
	Cum. Var. Exp	25.70176	49.25677	68.56660	77.52055
1	Day of 50% FLOWERING	0.03860	0.44256	0.36942	0.28700
2	Day to maturity	0.19988	-0.27914	-0.38417	-0.43618
3	Productive tiller per plant	-0.29372	-0.36637	0.25190	0.23345
4	Plant height	-0.32352	-0.42437	0.07713	-0.01739
5	Ear length per plant	-0.44118	0.17582	-0.23419	0.39959
6	No of spikelets per spike	-0.04668	0.17751	0.53878	-0.54957
7	No of grain per spike	-0.54275	0.07650	-0.09094	-0.12602
8	Weight of 1000 grains	0.21470	-0.47997	0.03563	0.28934
9	Chlorophyll content	-0.13886	-0.33310	0.49858	-0.06279
10	Grain yield per plant	-0.45918	0.04799	-0.20804	-0.32763

Table 2 : K - Clustering pattern of 40 Genotypes

Group	n	Within SS	Cluster Members
1	1	0.0000	VWFW2024.
2	7	2.9546	IC-252794, IC-443633, EC-27051, BW11, EC-478017, IC-542076, PBW-373(ch)
3	9	8.8416	IC-574476, IC-406521, IC-335683, RAJ164, UAS-415, EC-578134, IC-401927, EC-273814, K-1317(ch)
4	2	1.0849	IC-574388, IC-443766
5	11	30.1782	IC-116276, K9423, IC-290186, IC-122126, IC-113734, IC-527448, MIYCSN22, IC-554661, EC-464070, IC-566636, HD-2967(ch)
6	7	5.4832	EC-577050, IC-402042, IC-574387, EC-576930, IC-252429, K-1006(Ch), K-9107(ch)
7	3	5.7729	IC-531183, IC-290195, NINGMAI

Table 3 : K-Cluster mean for 7 cluster wheat genotypes

Cluster	Days to 50% Flowering	Day to maturity	Productive Tillers Per Plant	Plant Height	Ear length per plant (cm)	No of spikelet per plant	No of grain per spike	1000 Seed Weight (g)	Chlorophyll content	Grain Yield per plant
1 Cluster	76.333	111.333	9.667	101.000	14.500	19.667	88.000	37.333	48.200	13.000
2 Cluster	79.333	115.381	7.667	66.952	7.405	15.714	49.571	37.952	44.571	12.814
3 Cluster	79.519	113.889	10.074	80.074	9.148	17.185	53.148	32.889	44.137	13.056
4 Cluster	73.000	113.167	9.833	70.333	7.167	14.333	46.000	39.833	40.717	11.767
5 Cluster	75.667	113.727	13.606	92.303	9.591	16.788	51.970	38.621	45.164	13.718
6 Cluster	77.524	112.095	9.857	92.333	10.238	18.000	58.619	35.476	46.095	18.333
7 Cluster	80.000	112.778	9.222	72.000	10.556	21.000	78.111	32.056	45.322	13.133

Table 4: Mean performance of different forty wheat genotypes for ten characters

SN No.	Genotypes	Days to 50% flowering	days to maturity	productive tillers per plant	plant height(cm)	Ear length per plant (cm.)	No. of spikelets per plant	No of grain per spike	wt. of 1000 grains	Chlorophyll Content	grain yield per plant (gm)
1	IC-574388	73.33	113.67	10.00	72.67	8.33	16.00	47.00	38.67	41.50	12.17
2	EC-577050	78.67	112.67	8.67	98.00	10.00	19.67	46.67	31.33	45.00	9.60
3	IC-116276	73.33	112.67	15.33	99.33	10.17	14.33	52.67	39.33	47.37	10.53
4	K9423	71.00	114.67	14.33	85.67	9.17	17.67	68.67	37.33	43.00	13.83
5	VWFW2024	76.33	111.33	9.67	101.00	14.50	19.67	88.00	37.33	48.20	13.00
6	IC-290186	70.33	111.67	16.33	99.33	9.17	14.33	37.33	39.00	46.97	12.87
7	IC-402042	77.67	109.67	9.67	90.33	10.33	16.33	66.33	37.00	52.03	11.73
8	IC-531183	78.33	111.67	9.67	73.00	11.50	19.67	66.67	31.67	47.33	14.47
9	IC-122126	72.00	111.33	16.67	89.67	10.33	16.33	62.67	40.00	45.33	15.07
10	IC-252794	78.00	116.67	7.00	62.00	8.50	16.67	57.33	39.83	42.23	13.90
11	IC-574476	78.33	114.33	8.00	64.00	8.67	18.67	65.00	33.00	47.13	13.33
12	IC-443633	81.00	119.00	7.67	68.67	8.67	14.00	41.00	32.83	44.13	10.77
13	IC-406521	81.67	114.67	8.67	95.33	7.17	17.33	47.00	30.33	43.93	12.07
14	IC-335683	81.67	113.33	6.00	80.00	9.83	17.67	63.00	37.67	45.20	14.80
15	EC-27051	81.67	118.67	8.33	62.67	6.67	16.67	57.33	38.67	46.37	12.03
16	IC-113734	77.67	113.00	9.67	88.67	10.67	15.67	41.33	42.67	42.43	15.17
17	BW11	72.67	112.33	7.33	59.67	8.00	15.67	48.67	35.33	40.17	12.03
18	EC-478017	81.67	114.33	8.33	67.00	5.33	17.00	48.33	42.33	48.97	13.67
19	IC-443766	72.67	112.67	9.67	68.00	6.00	12.67	45.00	41.00	39.93	11.37
20	IC-542076	79.33	111.67	8.33	73.67	6.83	12.00	45.33	38.50	44.80	13.43
21	RAJ164	76.67	111.67	9.00	80.33	7.83	15.33	42.33	23.33	43.37	11.00
22	UAS-415	80.33	114.33	13.33	77.00	7.50	19.00	49.33	34.17	50.07	15.93
23	EC-578134	80.67	113.00	17.67	76.00	11.17	15.00	48.67	34.67	47.43	10.80
24	IC-401927	76.67	115.33	11.33	83.67	11.00	17.33	52.00	34.33	32.10	12.57
25	IC-527448	77.00	115.00	11.00	87.00	8.17	15.33	56.67	37.00	43.93	11.00
26	MIYCSN22	78.00	113.67	12.67	98.33	7.17	17.00	41.67	33.67	42.43	11.00
27	IC-574387	76.67	113.67	8.00	83.67	11.17	18.00	72.33	36.67	48.93	13.53
28	EC-576930	77.33	113.33	12.67	97.33	7.67	17.00	57.00	35.00	46.00	49.43
29	IC-554661	78.67	114.67	17.00	98.00	11.33	20.67	57.00	38.83	46.97	15.07
30	EC-273814	81.33	116.33	8.67	93.33	9.17	18.67	47.00	27.67	43.53	12.30
31	EC-464070	78.33	112.67	10.67	89.00	10.17	16.33	38.33	38.83	44.90	16.03
32	IC-290195	80.33	112.00	7.33	66.33	10.33	21.00	85.67	30.00	40.07	10.03
33	IC-252429	77.67	110.33	9.67	94.00	9.17	17.33	50.33	37.33	41.97	12.83
34	IC-566636	79.67	113.33	11.67	86.33	8.83	15.67	49.33	39.67	45.13	16.40
35	NINGMAI	81.33	114.67	10.67	76.67	9.83	22.33	82.00	34.50	48.57	14.90
36	K-1006(Ch)	77.33	114.33	9.67	88.00	10.67	16.67	66.67	36.83	40.30	12.77
37	K-1317(ch)	78.33	112.00	8.00	71.00	10.00	15.67	64.00	40.83	44.47	14.70
38	PBW-373(ch)	81.00	115.00	6.67	75.00	7.83	18.00	49.00	38.17	45.33	13.87
39	K-9107(ch)	77.33	110.67	10.67	95.00	12.67	21.00	51.00	34.17	48.43	18.43
40	HD-2967(ch)	76.33	118.33	14.33	94.00	10.33	21.33	66.00	38.50	48.33	13.93
	Mean	77.71	113.61	10.50	82.72	9.30	17.17	55.54	36.20	44.86	14.06
	C.V.	2.16	1.35	21.69	4.40	13.22	9.97	9.42	4.03	10.05	72.62
	C.D. 5%	10.21	5.86	5.39	34.72	6.76	5.57	17.19	22.81	1.89	1.05
	C.D. 1%	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.41

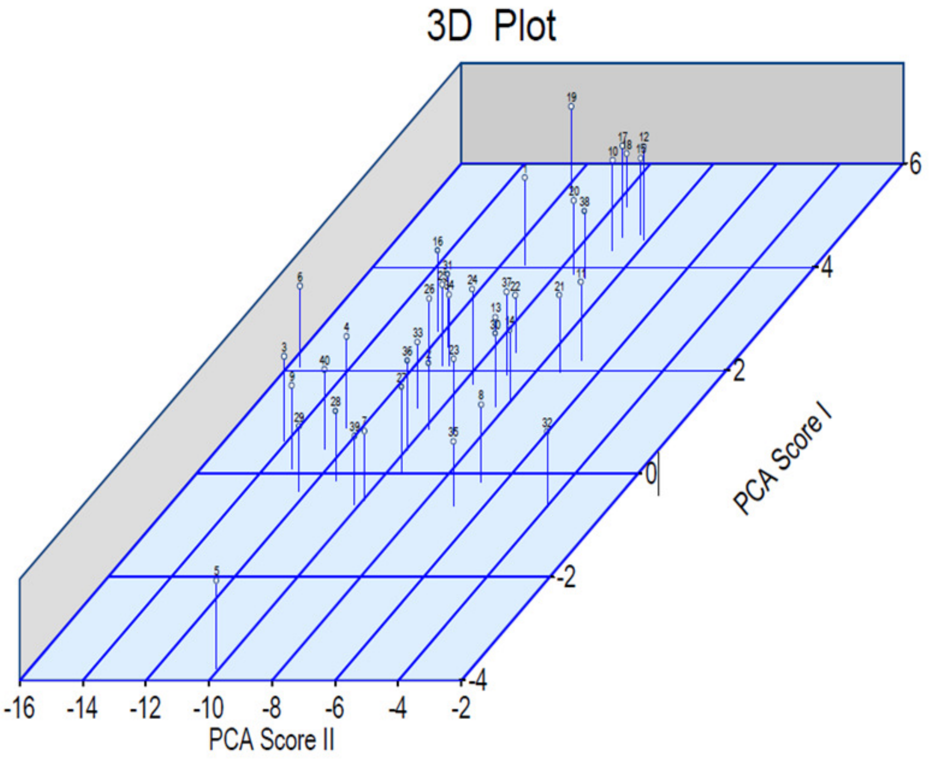

Fig. 1 : 3D plot of PCA Score I and II

Conflict of interest

All Authors have declared that no competing interests exist.

References

Curtis, T., & Halford, N. G. (2014). Food security: the challenge of increasing wheat yield and the importance of not compromising food safety. *Annals of applied biology*, **164**(3), 354-372.

Eivazi, A. R., Naghavi, M. R., Hajheidari, M., Pirseyedi, S. M., Ghaffari, M. R., Mohammadi, S. A., & Mardi, M. (2008). Assessing wheat (*Triticum aestivum* L.) genetic diversity using quality traits, amplified fragment length polymorphisms, simple sequence repeats and proteome analysis. *Annals of Applied Biology*, **152**(1), 81-91.

Hailegiorgis, D., Mesfin, M., & Genet, T. (2011). Genetic divergence analysis on some bread wheat genotypes grown in Ethiopia. *Journal of Central European Agriculture*, **12**(2), 0-0.

Harlan, J. R., & Zohary, D. (1966). Distribution of wild wheats and barley: the present distribution of wild forms may provide clues to the regions of early cereal domestication. *Science*, **153**(3740), 1074-1080.

Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A k-means clustering algorithm. *Journal of the royal statistical society. series c (applied statistics)*, **28**(1), 100-108.

Islam, M. R., Faruquei, M. A. B., Bhuiyan, M. A. R., Biswas, P. S., & Salam, M. A. (2004). Genetic diversity in irrigated rice. *Pakistan Journal of Biological Sciences*, **2**, 226-229.

Khan, M. A., Anjum, A., Bhat, M. A., Padder, B. A., Mir, Z. A., & Kamaluddin, M. (2015). Multivariate analysis for morphological diversity of bread wheat (*Triticum aestivum* L.) germplasm lines in Kashmir valley. *Journal of Science*, **5**(6), 372-376.

Kiszonas, A. M., & Morris, C. F. (2018). Wheat breeding for quality: A historical review. *Cereal Chemistry*, **95**(1), 17-34.

Lopes, M. S., El-Basyoni, I., Baenziger, P. S., Singh, S., Royo, C., Ozbek, K., ... & Vikram, P. (2015). Exploiting genetic diversity from landraces in wheat breeding for adaptation to climate change. *Journal of experimental botany*, **66**(12), 3477-3486.

Mellingers, J. S. (1972). Measures of Genetic Similarity and Genetic Distance. VII. *Studies in Genetics*, 145-153.

Mishra, C. N., Tiwari, V., Satish-Kumar, S. K., Gupta, V., Kumar, A., & Sharma, I. (2015). Genetic diversity and genotype by trait analysis for agromorphological and physiological traits of wheat (*Triticum aestivum* L.).

Mohammadi, S. A. (2002,). Statistical methods in genetics. In *Proceedings of the Sixth International Conference of Statistics. August* (pp. 26-28).

Rizwana Maqbool, R. M., Muhammad Sajjad, M. S., Ihsan Khaliq, I. K., Aziz-ur-Rehman, A. U. R., Khan, A. S., & Khan, S. H. (2010). Morphological diversity and traits association in bread wheat (*Triticum aestivum* L.).

Sharma, J. R. (2006). *Statistical and biometrical techniques in plant breeding*. New Age International.