International Journal
of Pure & Applied
Bioscience

Research Article

OPEN ACCESS

Genetic Diversity Based on Cluster and Principal Component Analyses for Yield and its Contributing Characters in RILs of Bread wheat (*Triticum aestivum* L.)

Pooja[1*], S. S. Dhanda[2] and R. S. Beniwal[3]

[1]Ph.D Student, [2]Principal Scientist, Wheat & Barley Section, Department of Genetics and Plant Breeding,
[3]Assistant Plant Pathologist, Department of Plant Pathology,
Chaudhary Charan Singh Haryana Agricultural University, Hisar-125004 Haryana, India
*Corresponding Author E-mail: deepooja16@gmail

ABSTRACT

An experiment was carried out to assess genetic diversity by cluster and principal component analysis (PCA) for yield and its ten contributing characters in 210 bread wheat recombinant inbred lines (RILs) at HAU Agricultural Research Farm during Rabi 2015-16 and 2016-17. The cluster analysis grouped all the 210 wheat RILs into three major clusters. Extreme divergence was observed among clusters. First cluster with seventeen lines had better yield potential and its contributing characters as compared with second and third cluster which had also six lines with desirable traits. The result of PCA revealed that all the 4 principal components (PC-I, PC-II, PC-III and PC-IV) contributed 78.34% of the total variability and accounted with values of 34.30, 25.87, 11.17 and 6.98 respectively. The first principal component had high positive component loading for all variables except 100-grain weight, harvest index and days to 50% heading. The result of present study could be exploited in planning and execution of future breeding improvement programme in wheat.

***Key words:** PCA, Cluster analysis, Diversity, Dendrogram, Eigen value, Bread wheat.*

INTRODUCTION

Sustainable production of food crops relies on germplasm improvement and genetic diversity. Genetic diversity analysis has opened new realm of crop improvement. Recent crop improvement efforts result in narrow genetic diversity of elite germplasm pool leads to problematic in breeding for biotic stresses, abiotic stresses as well as adaptation[1,2]. So, it is mandatory to identify genetically diversified lines for future wheat breeding. Information on genetic diversity among elite germplasm is beneficial for identifying lines for desirable traits[3] and estimating genetic relatedness among parents. Hybridization followed by subsequent selection is the important approach for wheat breeding programme. Joshi et al. reported that, for transgressive segregation, genetic variation between parents is necessary[4].

In order to maintain, evaluate and utilize germplasm effectively, it is important to investigate the extent of genetic diversity available. Higher the genetic distance between parents, the higher heterosis in progeny can be observed. Morphological traits have been successfully used for estimation of genetic diversity and cultivar development since they provide a simple way of quantifying genetic variation[5]. Therefore, concise information about nature and magnitude of genetic diversity present in wheat gene pool would be helpful regarding parents selection for deriving superior cultivars.

A number of methods are available for analysis of genetic diversity in germplasm accessions, breeding lines and populations. Multivariate analysis methods are also useful tool to access stability and can be used to identify groups with desirable traits for breeding [6]. Cluster method is an analysis (CA) that used dendrograms to display how various genotypes were differentiated. Diversity of tetraploid wheat germplasm grouped by cluster analyses (CA) and principal component analyses (PCA) explained the variation among genotypes[7,8]. As a multivariate statistical technique, the principal components analysis (PCA) has the ability to transform a number of possibly correlated variables into a smaller number of variables called principal components the principal components are linear transformations of the original variables and could be respective of a particular meaning [9]. This approach is very helpful in deciding which agronomic traits of crop contributing most to yield, subsequently, these agronomic traits should be emphasized in the breeding program.

MATERIAL AND METHODS

The current experiment was conducted in the Department of Genetics and Plant Breeding, Chaudhary Charan Singh Haryana Agricultural University, Hisar. Geographically, Hisar is located at latitude of 29.090N, longitude of 75.430E in western Haryana. It has an average elevation of 705ft above sea level. The region is part of the alluvial Ghaggar-Yumuna plain and its southern and western portions mark a gradual transition to the desert. Crop was grown in research area of Wheat Section, Department of Genetics and Plant Breeding, CCS HAU, Hisar, during the Rabi season of 2015-16 and 2016-17. Genetically pure and physically healthy seeds of 210 recombinant inbred lines (RILs) of diverse parents of WH542 X WH711 were obtained from Department of Genetics and Plant Breeding, CCS HAU Hisar. The experiment was carried out in Randomized Block Design (RBD) with three replications and plot was 3 m × 40cm size. The distance between row to row and plant to plant was maintained were 23 cm and 10 cm respectively. Recommended package of practices will be followed to raise the crop. Data will be recorded on the 5 plants/ replication/ genotype. Data were recorded on 11 characters, viz., yield/plant (g), number of tillers/ plant, number of grains/ear, number of spikelets/ear, 100-grain weight (g), Biological yield/plant (g), Harvest index (%), Plant height (cm), Ear length (cm), Days to 50% heading, Days to 75% maturity. The statistical tool INDOSTAT and XLSTAT was utilized for genetic divergence, principal component and cluster mean analysis including dendogram.

RESULTS AND DISCUSSION
Principal Component analysis (PCA)

The average data was analyzed by using Principal Component analysis. Principal component analysis reflects the importance of the largest contributor to the total variation at each axis of differentiation. The eigen values are often used to determine how many factors to retain. The sum of the eigen values is usually equal to the number of variables [10]. Three principal components showed more than one Eigen value and showed about 78.34% of variability. Table.1 PC-I showed 34.30%, PC-II 25.87%, PC-III 11.17% and PC-IV exhibited 6.98% variability among different traits of the RILs under experiment. Eigen value and variance associated with each principal component, decreased gradually and stopped at 0.64 and 6.98 respectively.

Table 1: Principal Components (PCs) analysis for yield and contributing traits in RILs of bread wheat

Traits	PC-I	PC-II	PC-III	PC-IV
Eigen value	3.16	2.39	1.03	0.64
Total Variance (%)	34.30	25.87	11.17	6.98
Cumulative %	34.30	60.17	71.35	78.34

The first PC was related to yield and yield contributing traits i.e. plant height, biological yield, ear length, number of tillers/plant, number of spikelets/ear, days to 75% maturity and number of grains/ear as cleared from the values of Table.2 for PC-I. Poor in 100-grain weight, harvest index and days to 50% heading. The results are agreed with Leilah & Al- Khateeb[11], Saifullah *et al*.,[12] who reported the more similar results for first principal component and reported the maximum variability in the data with respect to succeeding components. In second principal component exhibited positive effect for grain yield/plant, biological yield, number of tillers/plant and days to 75% maturity while the values of traits like number of grains/ear and days to 50% heading remained low (Table.2 PC-II). In third PC, number of spikelets/ear, 100-grain weight and grain yield/plant, ear length were the traits of more importance, the values for characteristics like days to 50% heading, number of tillers/plant and harvest index remained lower (Table. 2 PC-III). Similarly, in fourth principal component exhibited positive effect for grain yield/plant, harvest index and ear length while the values of traits like number of grains/ear and number of spikelets/ear remained low (Table.2 PC-IV).

In the present study, differentiation of the RILs into different clusters was because of relatively high contribution of few characters rather than small contribution from each character. The positive and negative loading shows the presence of positive and negative correlation trends between the components and the variables. Therefore, the above mentioned characters which load high positively or negatively contributed more to the diversity and they were the ones that most differentiated the clusters. Accordingly, the first principal component (PC-I) had high positive component loading from plant height, grain yield/plant, biological yield, ear length, number of tillers/plant, number of spikelets/spike, and number of grains/ear and exhibited high negative loading from harvest index, days to 50% heading and 100-grain weight. In the second principal component (PC-II) high positive component loading was observed from grain yield/plant, biological yield, number of tillers/plant while it was high negative from plant height, ear length, harvest index and number of spikelets/spike. The major contributing characters for the diversity in the third principal component (PC-III) were number of spikelets/ear, 100-grain weight, grain yield/plant exhibited high value for positive loadings and number of grains/ear, plant height, biological yield and days to 50% heading showed the high negative values for morphological traits. Similarly in the fourth principal component (PC-IV) high positive component loading was observed from grain yield/plant, harvest index, ear length, number of grains/ear and number of spikelets/ear while it was negative from biological yield, 100-grain weight, number of tillers/plant and plant height. Usually it is customary to choose one variable from these identified groups. Hence, for the first group plant height is best choice, which had the largest loading from component one, grain yield/plant for the second and fourth both and number of spikelets/ear for the third group.

Table 2: Factor loadings for yield contributing traits in RILs of bread wheat

Characters	PCA1	PCA2	PCA3	PCA4
Grain yield/plant	0.5530	0.5968	0.0691	0.4817
Number of tillers/plant	0.1916	0.1277	0.0437	-0.0885
100-grain weight	-0.0294	-0.0241	0.0698	-0.2499
Number of grains/ear	0.0208	0.0112	-0.9354	0.1352
Plant height	0.6870	-0.6855	-0.0459	-0.0375
Ear length	0.2347	-0.0155	0.0622	0.1426
Number of spikelets/ear	0.1023	-0.0489	0.2995	0.0227
Biological yield/ plant	0.2931	0.3791	-0.1300	-0.7065
Harvest index	-0.1342	-0.0804	0.0209	0.3605
Days to 50% heading	-0.1065	0.0123	-0.0149	-0.0149
Days to 75% maturity	0.0565	0.0614	0.0096	-0.1605
Expl.Var	3.61	2.39	1.03	0.64
Prp.Total	0.34.	0.25	0.11	0.06

Cluster mean analysis

The cluster analysis sequestrates genotypes into clusters which exhibit high homogeneity within a cluster and high heterogeneity between clusters[13]. According to Chahal and Gosal[14] characters with largest absolute value closer to unity within the first principal component influence the clustering more than those with lower absolute value closer to zero. Therefore, in the present study, differentiation of the genotypes into different clusters was because of relatively high contribution of few characters rather than small contribution from each character. All the three clusters were analyzed for mean and standard deviation (Table.4) Distribution pattern of all the RILs into three clusters showed the presence of considerable genetics diversity among the lines for most of the traits under consideration.

The characteristic feature of each cluster is discussed here below. The mean values for grain yield/plant (16.19), number of tillers/plant (10.40), 100-grain weight (4.35), plant height (81.91), number of spikelets/ear (21.33), biological yield (42.15), harvest index (24.51), days to 50% heading (99.7) and days to 75% maturity (141.52) was highest in cluster-I. Cluster-II exhibited the maximum mean value (58.89) for number of grains/ear. Average values for plant height (115.77) and ear length (10.05) was higher in cluster-III (Table 4).

Table 3: Grouping RILs using cluster analysis based on principal component analysis

I	102	P1,P2,1,2,4,5,6,7,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,40,41,43,50,52,54,55,60,61,65,70,71,72,73,74,75,76,77,81,82,83,86,88,92,93,96,97,98,100,104,106,107,108,109,110,111,112,113,114,115,117,118,119,120,122,123,124,125,127,130,132,133,134,135,138,140,141,142,149,151,152,155,172,173,193,207
II	46	3,8,44,46,51,53,57,58,62,63,64,79,80,84,85,87,89,90,95,99,105,128,137,143,144,147,153,159,162,163,164,169,170,181,185,188,189,190,191,195,198,199,205,206,209,210
III	64	37,38,39,42,45,47,48,49,56,59,66,67,68,69,78,91,94,101,102,103,116,121,126,129,131,136,139,145,146,148,150,154,156,157,158,160,161,165,166,167,168,171,174,175,176,177,178,179,180,182,183,184,186,187,192,194,196,197,200,201,202,203,204,208.

Table 4: Means, Standard deviations and Variances for clusters based on yield contributing traits

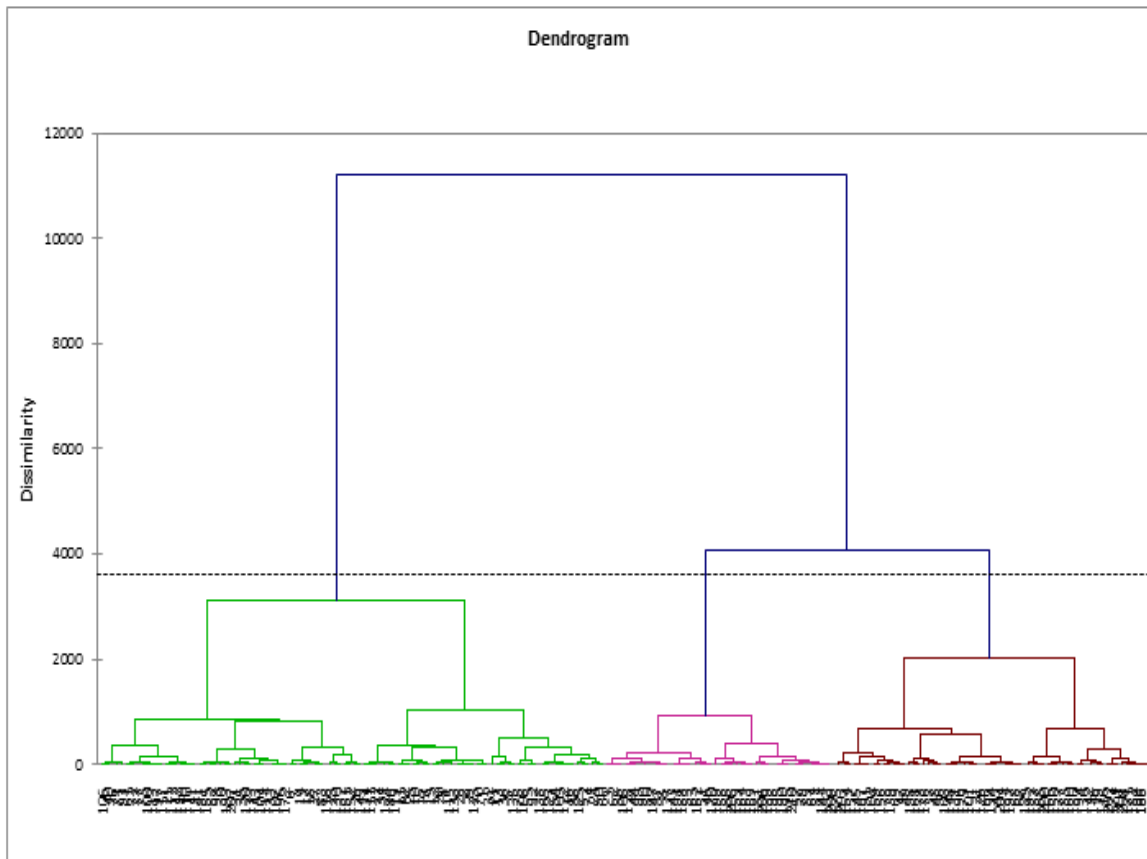| Characters	Cluster-I		Cluster-II		Cluster-III	
	Mean	SD	Mean	SD	Mean	SD
Grain yield/plant	16.19	1.17	12.60	1.19	10.27	1.60
Number of tillers/plant	10.40	1.42	9.66	1.48	9.65	1.65
100-grain weight	4.35	1.21	4.30	1.31	4.31	1.57
Number of grains/ear	52.98	1.39	58.89	1.49	49.90	1.66
Plant height	81.91	1.56	77.66	1.34	81.55	1.96
Ear length	10.05	1.61	9.60	1.54	10.20	1.45
Number of spikelets/ear	21.33	1.51	21.09	1.51	21.14	1.48
Biological yield/ plant	42.15	1.64	33.38	1.57	26.59	1.43
Harvest index	24.51	1.87	21.36	1.54	19.21	1.60
Days to 50% heading	99.97	1.76	100.54	1.59	100.61	1.56
Days to 75% maturity	141.52	1.58	142.78	1.72	142.17	1.88

Fig. 1: Dendrogram showing relationship among different RILs of wheat

Dendrogram showed that recombinant inbred line have more genetic distance among them. The results of the genetic distance has shown that there is a room for the genetic improvement of bread wheat varieties and the information generated can be used to plan wide crosses, to exploit genetic diversity and maximize the expression of hetrosis.
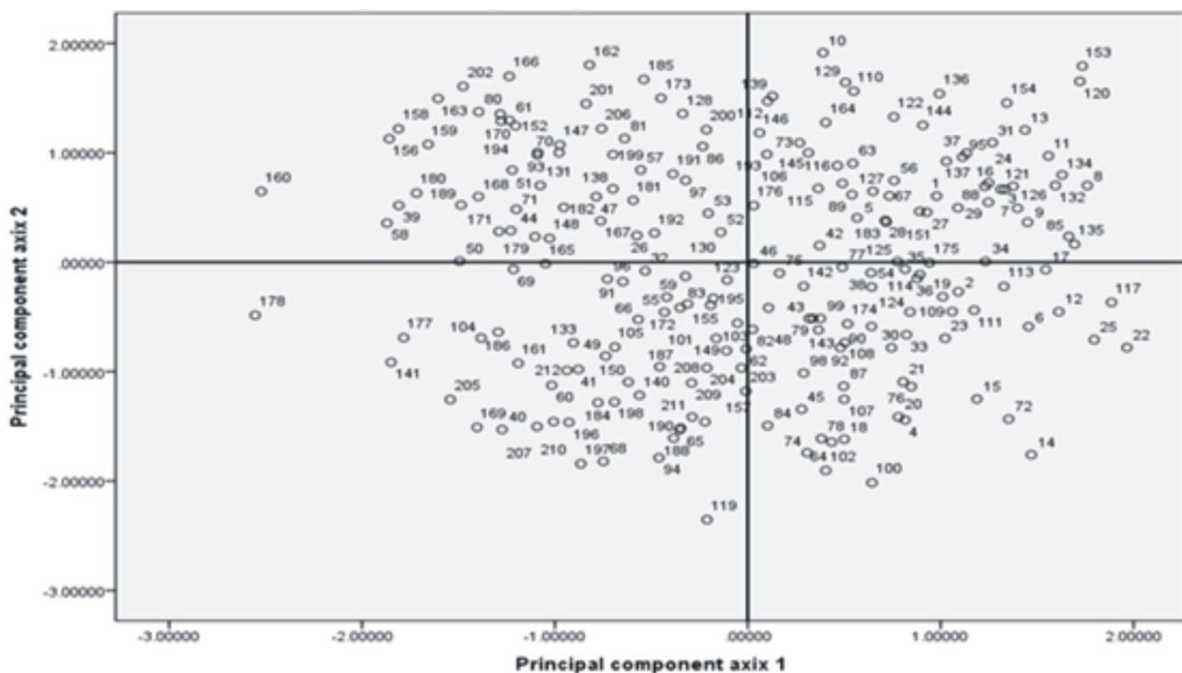

Fig. 2: Scatter diagram of wheat 210 RILs for first two PCs

CONCLUSION

The present research provided significant information helpful in genetic improvement of bread wheat. Considering the four main components, explained 87.34% of total variations in the data. However results showed that cluster analysis based on PCA is more precise indicator of differences among wheat RILs than cluster analysis (not based on PCA). Evaluation of genetic diversity can be useful for the selection of the most efficient lines for the production with higher uniformity may guarantee the production of enough food for the world increasing population.

Acknowledgement

REFERENCES

1. Rao, R.V., and Hodgkin, T., Genetic diversity, conservation and utilization of plant genetic resources. *Plant Cell, Tissue and Organ Culture,* **68**: 1-19 (2002).
2. Zhang, P., Dreisigacker, S.A.E., Melchinger, J.C., Reif, A., Mujeeb, K.M., Van, G., Hoisington, D. and Warburton, M.L., Quantifying novel sequence variation and selective advantage in synthetic hexaploid wheats and their backcross-derived lines using SSR markers. *Molecular Breeding,* **15:** 1-10 (2005).
3. Ali, Y., Atta, B.M., Akhter, J., Monneveux, P. and Lateef, Z., Genetic variability, association and diversity studies in wheat (*Triticum aestivum* L.) germplasm. *Pakistan Journal of Botany,* **40**: 2087-2097 (2008).
4. Joshi, B.K., Mudwari, A., Bhatta, M.R., Ferrara, G.O., Genetic diversity in Nepalese wheat cultivars based on agro-morphological traits and coefficients of parentage. *Nepal Agriculture Research Journal*, **5:** 7-17 (2004).
5. Fufa, H., Baenizger, P.S., Beecher, B.S., Dweikat, I., Graybosch, R.A. and Eskridge, K.M., Comparison of phenotypic and molecular-based classifications of hard red winter wheat cultivars. *Euphytica*, **145:** 133-146 (2005).
6. Lin, C.S., Binns, M.R. and Lefkovitch, L.P., Stability analysis: Where do we stand? *Crop Science*, **26:** 894-900 (1986).
7. Hailu, F., Merker, A., Harjit-Singh, Belay, G. and Johansson, E., Multivariate analysis of diversity of tetraploid wheat germplasm from Ethiopia. *Genetic Resources and Crop Evolution*, **53**:1089-1098 (2006).
8. Skrbic, B. and Onjia, A., Multivariate analyses of microelement contents in wheat cultivated in Serbia. *Food Control*, **18 (4):** 338- 345 (2002).
9. Jobson, D., Confidence Regions for the Mean Variance Set: An Alternative Approach to Estimation Risk. *Review of Quantitative Finance and Accounting,* **1:** 235–257 (1991).
10. Sharma, J.R., Statistical and Biometrical Techniques in Plant Breeding. *New Age International (P) Limited Publishers, New Delhi*, 432 (1998).
11. Leilah, A.A. and Al-Khateeb, SA., Statistical analysis of wheat yield under drought conditions. *Journal of Arid Environments*, **61(3)**: 483-496 (2005).
12. Saif-Ullah, A., Minhas, N.M., Hamdani, A., Shakir, A., Zubair, M. and Ahmad, Z., Multivariate analysis of genetic divergence in wheat (*Triticum aestivum*) germplasm. *Pakistan Journal Botany*, **45(5):** 1643-1648 (2013).
13. Jaynes, D.B., Kaspar, T.C., Colvin, T.S. and James, D.E., Cluster analysis of spatiotemporal corn yield patterns in an Iowa field. *Agronomy Journal,* **95:** 574–586 (2003).
14. Chahal, G.S. and Gosal, S.S., Principles and Procedures of Plant Breeding, Biotechnology and Conventional Approaches. *Narosa Publishing House*, New Delhi (2002).