Phenotypic Diversity of Pea Genotypes (*Pisum sativum* L.) Based on Multivariate Analysis

Shalini Singh, V. Rakesh Sharma¹, Vinay Kumar Reddy Nannuru², B. Singh, Mukesh Kumar **10.18805/LR-4165**

ABSTRACT

The present study was conducted to identify the nature and magnitude of genetic divergence among fifty five genotypes based on phenotypical traits using the multivariate analysis. Based on cluster analysis, the genotypes were best fitted into six clusters. The maximum and minimum genotypes grouped in cluster III (14) and minimum were in Cluster IV (4) respectively. The maximum intracluster distance was shown by cluster I (D^2 2.543) indicating maximum difference among the genotypes within and the minimum value was shown by cluster II (D² =1.827). Maximum value of inter-cluster distance (D²=6.471) was recorded between cluster I and cluster IV revealing that the genotypes of these clusters were highly diverse from others and can be used as divergent parents for hybridization and selection. Thus, for getting high heterosis for recovering transgressive segregants, genotypes from cluster I and IV can be used as distant parents in any breeding programme for successful pea improvement. The Optimum number of K-means clusters formed were four. The result of PCA revealed that all the four principal components (PC-I, PC-II, PC-III and PC-IV) contributed 86.7% of the total variability. The results of present study could be exploited in the future genetic improvement programme of pea genotypes in Uttar Pradesh region.

Key words: Genetic diversity, Hierarchical clustering, K-means clustering, Pea.

INTRODUCTION

Pea (Pisum sativum L.) 2n=2x=14 is one of the world's oldest crop which belongs to the family Leguminosae grown in all the temperate countries and in the most tropical highlands. It has been grown for several thousand years in India and is a crop native to Syria, Turkey, Israel and Ethiopia, for its versatile uses as pulses and livestock feed (Choudhury *et al*., 2007). It is grown for its green tender pods, dried seeds, canned, frozen or dehydrated form (Santalla *et al*., 2001).

India shares 21 per cent of the world's pea production with an area of 5.43 lakh hectare and the production of 54.32 lakh tons (NHB, 2018). Most of the production is produced in Uttar Pradesh, Madhya Pradesh, Bihar, Assam and Orissa, and the total area is about 95 per cent. Uttar Pradesh is a major pea producing state in India producing about 60 per cent of the country's produce. They are rich source of phytonutrients, minerals, vitamins and antioxidants and is known for its superior quality protein like high levels of lysine making it an appropriate dietary complement to the cereals (Dhama *et al*., 2010). Being a short duration crop it is highly utilized for crop rotation and also have an important role in the modern agricultural systems as it is capable to enhance the soil structure and provides breaks for the disease control (Martin *et al*., 2008).

Substitution of landraces and traditional pea accessions by the modern cultivars is widespread and consequently the genetic variability loss is reduced, in particular replacement with cultivars characterized by superior tolerance for biotic and abiotic stress (Handerson *et al*., 2014). In the process of genetic improvement of any crop,

Department of Horticulture, Sardar Vallabhbhai Patel University of Agriculture and Technology, Meerut-250 110, Uttar Pradesh, India. ¹CSIR-National Botanical Research Institute, Rana Pratap Marg, Lucknow-226 001, Uttar Pradesh, India.

²Crop Biodiversity and Breeding Informatics, University of Hohenheim, Germany.

Corresponding Author: Shalini Singh, Department of Horticulture, Sardar Vallabhbhai Patel University of Agriculture and Technology, Meerut-250 110, Uttar Pradesh, India. Email: shalini.singh27292@gmail.com

How to cite this article: Singh, S., Sharma, V.R., Nannuru, V.K.R., Singh, B. and Kumar, M. (2021). Phenotypic Diversity of Pea Genotypes (*Pisum sativum* L.) Based on Multivariate Analysis. Legume Research. 44(8): 875-881. DOI: 10.18805/LR-4165.

Submitted: 11-05-2019 **Accepted:** 30-08-2019 **Online:** 09-11-2019

genetic diversity among germplasm plays a major role, since it opens the way to determine the most divergent parents based on the contribution of different qualitative and quantitative traits, for further utilization in any hybridization programme. Therefore, the exploration of genetic diversity in the available germplasm is a pre-requisite in a breeding programme for effective selection of the superior genotypes. A plant breeder has to identify the source of favorable genes to incorporate them into the breeding populations and select for a combination of desirable traits that might result in the isolation of productive genotypes and cultivars. Thus, present study is undertaken to understand the magnitude of genetic divergence for identifying more diverse parents for pea genetic improvement.

MATERIALS AND METHODS

The present investigation consisting of fifty five genotypes of pea was carried out at Horticultural Research Centre, SVPUA&T, Meerut, U.P. (29°01 latitude N and 77°43 longitude E and 219.75 MSL) during *Rabi* season 2015-16 to evaluate the amount of genetic diversity and relatedness among the genotypes (Table 1). The experimental design was RBD with three replications. In each replication the seeds were sown in a plot of 6.0 $m \times 1.0$ m in which row to row and plant to plant spacing was 60 cm and 10 cm respectively. All the recommended agro-practices were followed to ensure a healthy crop growth and development. Observation were recorded on five competitive plants situated under the same field condition for eleven morphological quantitative traits *viz*., days to 50 per cent flowering, plant height, number of first fruiting node, length of first fruiting node, number of pods per plant, length of pod, width of pod, number of seeds per pods, green pod yield per plant, seed yield per plant and shell weight per plant. The magnitude of genetic diversity among fifty five pea genotypes was determined by using D**²**Mahalanobis genetic distance statistics (Mahalanobis, 1936). Hierarchical clustering using Tocher's method, as described by Rao (1952) was followed for the grouping of genotypes into distinct clusters. Partition clustering using K-means as suggested by Charrad *et al*. (2014) and principal component analysis were done using R software (Venables and Ripley, 2002).

RESULTS AND DISCUSSION

The data obtained from the observations recorded on eleven morphological quantitative traits were subjected to the statistical scrutiny. It was evident from the analysis of variance that mean sum of squares due to 55 genotypes were highly significant for all the traits except days to 50 per cent flowering (Table 2), giving the clear picture of presence of wide spectrum of variability among the genotypes. These results were in agreement with the findings of Lal *et al*. (2011); Supe *et al*. (2013); Georgieva *et al*. (2016) and Kumar and Kumar (2016). Although the analysis of variance revealed sufficient variability among the genotypes, but the extent of genetic diversity present among the genotypes could not be explained, therefore, cluster analysis was performed to quantify the genetic divergence between any two genotypes or group of genotypes.

Based on the relative magnitude of their Mahalanobis $D²$ values using Torcher's method, all the 55 genotypes of pea under study were grouped into six clusters. The clustering patterns of pea genotypes into six clusters are presented in Table 3. Maximum number of genotypes (14) was grouped in cluster III namely: VRP-13, VRP-375, VRP-324, VRP-311, VRP-176, VRP-327, VRP-276, VRP-82, VRP-248, KS-228, DPP-94/8-06, Kashi Uday, Kashi Samridhi and Kashi Nandini. Whereas, cluster I and cluster V both contained ten genotypes each where, cluster I comprises of genotypes namely: VRP-3, VRP-228, VRP-320, VRP-22, VRP-122, VRP-383, VRP-402, VRP-382, VRP-145

and Kashi Mukti; and cluster V consisted of genotypes namely: VRP-69, VRP-313, VRP-73, VRP-321, VRP-16, VRP-284, VRP-223, VRP-343, VRPM-15 and Kashi Shakti. Furthermore, cluster II comprises of nine genotypes namely: VRP-26, VRP-273, VRP-107, VRP-156, VRP-174, VRP-49, VRP-131, EC-8724 and MO-23; followed by eight genotypes that were arranged in cluster VI namely: VRP-194, VRP-115, VRP-355, VRP-65, VRP-64, VP-233, EC-97280 and EC-8372. Cluster IV consisted with minimum of four genotypes namely: VRP-222, VRP-95, EC-71944 and MO-19. No parallelism was shown by the grouping pattern of the genotypes between the genetic diversity and geographical origin of genotypes. Similar confirmations were also reported by the findings of Singh *et al.* (2007); Dhama *et al*. (2009); Katiyar and Dixit (2009); Yadav *et al*. (2009); Devi *et al*. (2010); Shrivastava *et al.* (2012); Supe *et al*. (2013) and Kumar and Kumar (2016).

The average intra and inter-cluster D^2 values with their corresponding intra and inter-cluster distance are presented in Table 4. The inter-cluster distances were greater than intracluster distances, which indicated the presence of considerable amount of genetic diversity among the genotypes studied. The greater the magnitude of intra and inter cluster distance the higher the variability among the cluster and within the cluster and *vice versa*. The results are in concurrence with the findings of Kumar *et al*. (2006); Singh *et al*. (2007); Singh and Mishra (2008); Katiyar and Dixit (2009); Sen and De (2017). The least value of intra cluster distance was found in cluster II (D^2 = 1.827) indicating the presence of less heterogeneous genotypes grouped in this cluster. Whereas, maximum value of intra-cluster distance was observed in cluster I (D^2 = 2.543) revealing the existence of maximum differences among the genotypes falling in this cluster, followed by divergence ($D^2 = 2.352$) for cluster VI, cluster III with ($D^2 = 2.340$), cluster V ($D^2 = 2.158$) and cluster IV ($D^2 = 2.041$). Hence, selection within these clusters may be exercised based on the highest area of desirable traits. In any breeding programme where the nature of crosses is to be evaluated, choice of diverse parents is of paramount importance as they produce superior off-springs in the segregating generation than the closely related ones. The inter-cluster distance (D^2) being the main criterion for selection of genotypes was also worked-out as crossing of genotypes within the same cluster would not produce superior off-springs. A range of 2.462 to 6.471 was observed when inter-cluster D^2 values were used to study the diversity

Table 2: Analysis of variance for 11 characters in pea (*Pisum sativum* L.).

DF	DF (50%)	PH	NFFN	LFFN	NPP	LP	WP	NSP	GPY	SW	SY
2	0.12	1.28	0.04	0.78	0.07	0.06	0.004	0.02	0.00	0.01	0.06
54	16.54	1680.55**	$12.13**$	135.01**	69.08**	$2.72**$	$0.087**$	$2.17**$	508.53**	145.58**	142.64**
	0.28	8.49	0.31	3.07	0.27	0.01	0.005	0.03	2.91	0.99	1.84
		108									

**Significant at 1% level. DF- Degree of freedom, DF(50%)- Days to 50% flowering, PH- Plant height(cm), NFFN- Number of first fruiting node, LFFN- Length of first fruiting node (cm), NPP-Number of pods per plant, LP- Length of pod (cm), WP- Width of pod (cm), NSP-Number of seeds per pod, GPY- Green pod yield per plant (g), SW- Shell weight per plant (g), SY- Seed yield per plant (g).

Table 3: Clustering pattern of 55 pea genotypes on the basis of their mahalanobis genetic divergence using tocher's methods.

Clusters	Number of	Name of genotypes included						
	genotypes							
	10	VRP-3, VRP-228, VRP-320, VRP-22, VRP-122, VRP-383, VRP-402, VRP-382, VRP-145 and Kashi Mukti.						
\mathbf{H}	9	VRP-26, VRP-273, VRP-107, VRP-156, VRP-174, VRP-49, VRP-131, EC-8724 and MO-23.						
Ш	14	VRP-13, VRP-375, VRP-324, VRP-311, VRP-176, VRP-327, VRP-276, VRP-82, VRP-248, KS-228, DPP-						
		94/8-06, Kashi Uday, Kashi Sammridhi and Kashi Nandini						
IV	4	VRP-222, VRP-95, EC-71944 and MO-19						
V	10	VRP-69, VRP-313, VRP-73, VRP-321, VRP-16, VRP-284, VRP-223, VRP-343, VRPM-15 and Kashi Shakti						
VI	8	VRP-194, VRP-115, VRP-355, VRP-65, VRP-64, VP-233, EC-97280 and EC-8372.						

(Values are square root of D^2 value).

among the clusters. The minimum value of inter-cluster distance ($D^2 = 2.462$) was found between cluster II and III indicating close relationship and similarity for most traits among the genotypes included in these clusters. Whereas, cluster I and IV showed maximum value of inter-cluster distance ($D^2 = 6.471$), followed by cluster IV and V ($D^2 = 4.700$) and cluster I and III (D^2 = 4.626) indicating that the genotypes included in these clusters are not so closely related showing good amount of diversity. Hence, these genetically diverse genotypes can be used as promising parents for hybridization. These results are corroborated with the findings of Kumar *et al*. (2007); Singh *et al*. (2007); Devi *et al*. (2010) and Shrivastava *et al*. (2012) as they also gave similar conclusion.

Diversity among the genotypes was also estimated based on the considerable amount of variation in cluster means for different character. Different clusters exhibited distinct mean values for almost all the sixteen characters which reflect the genetic differences between the clusters (Table 5). It is evident from the cluster mean table that the genotypes in cluster I had highest mean values for number of pods per plant, green pod yield per plant, shell weight per plant and seed yield per plant. Whereas, the genotypes of cluster IV showed the maximum mean for days to 50 per cent flowering, plant height, length of first fruiting node and number of seeds per pod. Comparative assessment of cluster means showed that for improving specific characters, the genotypes should be selected from the cluster having high mean value for that particular character. This comparison indicates that clusters I and IV had better cluster means for most of the characters, therefore, these clusters might be considered better for selecting genotypes as divergent parents. The similar results are exhibited with the findings of Kumar *et al*. (2006); Devi *et al*. (2010) and Shrivastava *et al.* (2012).

The principal component analysis (PCA) is one of a series of techniques for collecting high-dimensional information and using the dependence between the variables in a more tractable form without any loss of information. It represents the major contributor to the total difference in each differentiation axis. Based on the analysis the first five principal components having eigen values greater than one contributed 86.70 per cent of the total variability of 55 pea germplasm (Table 6 and Fig 1). Proportion of variance for the first 4 components were 35, 19.2, 14.4, 10.9 per cent respectively. PC-I showed positive association towards days to 50 per cent flowering, plant height, number of first fruiting node and length of first fruiting

Table 5: Cluster wise mean values of 11 morphological traits in pea (*Pisum sativum* L.).

Clusters	DF (50%)	PH	NFFN	LFFN	NPP	LР.	WP	NSP	GPY	SW	SY
	62.43	60.97*	$7.10*$	$22.47*$	15.52**	7.60	1.33	5.33	47.52**	23.54**	$24.57**$
\mathbf{H}	63.70	64.89	8.00	25.63	7.84	6.52	$1.15*$	$4.87*$	19.01	9.92	9.25
Ш	62.19	69.45	8.66	28.14	$5.78*$	7.50	$1.47**$	5.47	17.60*	8.90	8.50
IV	66.25**	139.61**	8.27	35.01**	6.05	$6.48*$	1.22	$6.27**$	15.29	$8.39*$	$6.93*$
V	62.07*	73.40	10.35	34.54	9.73	$8.06**$	1.34	5.97	34.42	17.82	16.82
VI	65.08	79.18	$10.72**$	34.67	12.26	6.61	1.19	5.06	26.56	14.84	11.47

DF(50%)- Days to 50% flowering, PH- Plant height(cm), NFFN- Number of first fruiting node, LFFN- Length of first fruiting node (cm), NPP- Number of pods per plant, LP- Length of pod (cm), WP- Width of pod (cm), NSP- Number of seeds per pod, GPY- Green pod yield per plant (g), SW- Shell weight per plant (g), SY- Seed yield per plant (g).

Table 6: Factor loadings for yield contributing traits in Pea (*Pisum sativum* L.).

Traits	PC ₁	PC ₂	PC ₃	PC4	PC ₅	PC ₆	PC7	PC8
DF 50	0.037	-0.005	0.045	-0.350	0.480	-0.776	0.180	0.062
PH	0.946	-0.280	0.149	0.036	-0.031	0.013	-0.041	0.013
NFFN	0.008	0.001	-0.274	-0.084	0.145	-0.077	-0.920	0.116
LFFN	0.148	-0.041	-0.944	0.071	0.061	0.015	0.273	-0.017
NPP	-0.045	-0.257	0.018	-0.578	0.461	0.554	0.068	-0.234
LP	-0.011	-0.013	-0.009	0.162	-0.155	-0.135	0.058	-0.530
WP	-0.001	0.001	0.000	0.013	-0.014	-0.001	-0.015	-0.002
NSP	0.008	-0.005	-0.001	0.112	0.091	-0.144	-0.185	-0.795
GPY	-0.227	-0.750	-0.013	0.144	-0.080	-0.099	-0.003	0.038
SW	-0.105	-0.394	-0.065	-0.431	-0.546	-0.187	-0.042	-0.009
SY	-0.128	-0.369	0.076	0.533	0.440	0.022	0.001	0.117
Eigen values	2.963	2.453	1.760	1.493	1.020	0.823	0.684	0.362
Proportion of variance	0.350	0.192	0.144	0.109	0.072	0.062	0.042	0.012
Cumulative proportion (%)	35.00	54.20	68.70	79.50	86.70	92.90	97.10	98.30

DF(50%)- Days to 50% flowering, PH- Plant height(cm), NFFN- Number of first fruiting node, LFFN- Length of first fruiting node (cm), NPP- Number of pods per plant, LP- Length of pod (cm), WP-Width of pod (cm), NSP- Number of seeds per pod, GPY- Green pod yield per plant (g), SW- Shell weight per plant (g), SY- Seed yield per plant (g).

Fig 1: Screen plot showing variance and along with principal components.

Fig 2: Plot showing optimum K value of clustering for K-means clustering.

Fig 4: Scattered plot of pea genotypes using first two principal components.

Fig 5: Dendrograph showing relationship among pea genotypes using K-means clustering and hierarchical clustering based on mahalanobis genetic distance.

node. Whereas, PC-II positively associated with number of first fruiting node and width of the pod. Similar findings were reported by Maqbool *et al*. (2010) and Baranwal *et al*. (2013). For first two principal components, explained variation among all the pea genotypes was graphically represented in scattered plot (Fig 4).

Partitioning clustering was also performed by K-means cluster analysis based Mahalanobis genetic distance. Nbclust: R package used to obtain optimum K value (K=4) for clustering 55 pea genotypes collected from the two different locations Fig 2. Clustering pattern pea genotypes using K means clustering were presented in Fig 3. Moreover, the dendrograph showing both hierarchical clustering and K-means (K=4) based partitioning clustering presented in Fig 5. This result was in accordance to Charrad *et al*. (2014).

CONCLUSION

Based on Mahalanobis D^2 analysis it can be concluded that the pea germplasm in the present study can be successfully used for planning future breeding programmes. The intercrossing of genotypes showing the greater genetic divergence for most of the characters studied should result in superior heterotic crosses and also, generate valuable segregants in the later generations. It is expected that better performing varieties could be generated to increase productivity in field pea. Therefore, from the present study genotypes of cluster I and cluster IV based on their high values for inter cluster distance, cluster means and K clustering can be hybridized as the potential parents to produce superior off-springs in the segregating generations and to improve pea productivity.

REFERENCES

- Baranwal, D.K., Mishra, V.K., Singh, T. (2013). Genetic diversity based on cluster and principal component analysis for yield and its contributing characters in wheat (*Triticum aestivum* L.). Madras Agricultural Journal. 100 (4-6): 320- 323.
- Charrad, M., Ghazzali, N., Boiteau, V., Niknafs, A. (2014). NbClust: An R package for determining the relevant number of clusters in a data set. Journal of Statistical Software. 61(6): 1-36. URL http://www.jstatsoft.org/v61/i06/.
- Choudhury, P.R., Tanveer, H., Dixit, G.P. (2007). Identification and detection of genetic relatedness among important varieties of pea (*Pisum sativum* L.) grown in India. Genetica. 130: 183. https://doi.org/10.1007/s10709-006-9005-9.
- Devi, P.O., Pant, S.C., Rawat, S.S., Rana, D.K., Singh, N.I.K. (2010). Correlation coefficient and genetic divergence analysis in pea. Indian Journal of Horticulture. 67: 160-165.
- Dhama, S.K., Tyagi, N.K., Sirohi., S.P.S. (2009). Studies on genetic diversity under varying environments in pea (*Pisum sativum* L.). Legume Research. 32: 173-179.
- Dhama, S.K., Tyagi, N.K., Singh, P.B. (2010). Interrelationship and path analysis for seed yield and its component characters under eight environments in pea (*Pisum sativum* L.). Legume Research. 33(2): 87-94.
- Georgieva, N., Nikolova, I., Kosev, V. (2016). Evaluation of genetic divergence and heritability in pea (*Pisum sativum* L.). Journal of BioScience and Biotechnology. 5(1): 61-67.
- Handerson, C., Noren, S.K., Wricha, T., Meetei, N.T., Khanna, V.K., Pattanayak, A., Datt, S., Choudhury, P.R., Kumar, M. (2014). Assessment of genetic diversity in pea (*Pisum sativum* L.) using morphological and molecular markers. Indian Journal of Genetics. 74(2): 205-212.
- Katiyar, P.K. and Dixit, G.P. (2009). Multivariate analysis for genetic divergence in field pea [*Pisum sativum* (L.) var. *arvence*] germplasm. Indian Journal of Agricultural Sciences. 79: 181-183.
- Kumar, R. and Kumar, M. (2016). Estimation of genetic divergence in garden pea (*Pisum sativum var hortense* L.) germplasm to facilitate the selection of potential parents for hybridization programme. Legume Research. 39(5): 709-712.
- Kumar, R., Dhari, R., Kumar, R. (2006). Divergence studies in pea germplasm (*Pisum sativum* L.). National Journal of Plant Improvement. 8(2):122-124.
- Kumar, R., Dhari, R., Kumar, R., Malik, B.P.S. (2007). Assessment of morphological variability and genetic diversity in pea germplasm (*Pisum sativum* L.). National Journal of Plant Improvement. 9(1): 5-8.
- Lal, G.M., Meena, M.L., Chandra, K., Singh, C.M. (2011). Assessment of genetic variability and interrelation between yield and its contributing components in field pea (*Pisum sativum* L.). Environment and Ecology. 29: 1235-1239.
- Mahalanobis, P.C. (1936). On the Generalized Distance in Statistics. Proceedings of National Institute of Sciences (India). 2(1): 49-55.
- Maqbool, R., Sajjad, M., Khaliq, I., Rehman, A.U., Khan, A.S., Khan, S.H. (2010). Morphological diversity and traits association in bread wheat (*Triticum aestivum* L.). American-European Journal of Agriculture and Environment Science. 8: 216-224.
- Martin, R.J., Wilson, D.R., Butler, R.C., Riddle, M.U., Russell, A.C., Catherwood, D., Beare, M.H., *et al* (2008). Improving field pea yields on farm in canterbury. Proceedings of the Agronomy. Society of New Zealand. 38(4): 41-50.
- NHB, (2018). http://nhb.gov.in/statistics/State _Level/-2017-18 (3rd%20est)%20- %20data.pdf
- Rao, C.R. (1952). Advance Statistical Methods in Biometrical Research. 1st Ed. John Wiley and Sons, New York.
- Santalla, M., Amurrio, J.M., De Ron, A.M. (2001). Food and feed potential breeding value of green, dry and vegetable pea germplasm. Canadian Journal of Plant Science*.* 81(4): 601-610.
- Sen, M. and De, D.K. (2017). Genetic divergence in mung bean. Legume Research. 40(1): 16-21.
- Shrivastava, V., Lal, G.M., Chandra, K., Singh, C.M. (2012). Estimation of genetic distance in field pea (*Pisum sativum* L.). Environment and Ecology. 30(1): 136-138.
- Singh, D. and Mishra, V.K. (2008). Studies o genetic divergence in pea (*Pisum sativum* L.). Agricultural Science Digest. 20(1): 77-78.
- Singh, I., Singh, P., Sandhu J.S. (2007). Genetic divergence and association studies in field pea (*Pisum sativum* L.). Crop Improvement. 34(2): 179-182.
- Supe, V.S., Patil, S.D., Joshi, V.R. (2013). Genetic variability and diversity studies in pea (*Pisum sativum* L.). Journal of Agriculture Research and Technology. 38(2): 196-200.
- Venables, W.N. and Ripley, B. D. (2002). Modern Applied Statistics with S, Springer-Verlag.
- Yadav, R., Srivastava, Kant, R., Singh, R. (2009). Studies on genetic divergence in field pea [*Pisum sativum* (L.) var. *arvenc*]. Legume Research. 32: 121-124.