PLANT SCIENCE TODAY
Vol x(x): xx–xx
https://doi.org/10.14719/pst.9079

eISSN 2348-1900

RESEARCH ARTICLE

Genetic divergence and principal component analysis for yield and yield components of pigeon pea [*Cajanus cajan* (L.) Millsp.] germplasm

Anuj Kumar[1], Nitin Kumar[2], Vijay Sharma[2], Hoshiyar Singh Negi[3], Shalesh Kumar Singh[2], Kamaluddin[2]* & Vaishali Gangwar[4]

[1]Department of Genetics and Plant Breeding, Sardar Vallabhbhai Patel University of Agriculture and Technology, Meerut 250 110, Uttar Pradesh, India
[2]Department of Genetics and Plant Breeding, Banda University of Agriculture and Technology, Banda 210 001, Uttar Pradesh, India
[3]Department of Plant Pathology, College of Horticulture, Maharana Pratap Horticultural University, Anjanthali, Karnal 132 117, Uttar Pradesh, India
[4]Deparment of Crop Physiology, Banda University of Agriculture and Technology, Banda 210 001, Uttar Pradesh, India

*Correspondence email - kamaluddinpbg@gmail.com

Received: 25 April 2025; Accepted: 26 July 2025; Available online: Version 1.0: 11 September 2025

Cite this article: Anuj K, Nitin K, Vijay S, Hoshiyar SN, Shalesh KS, Kamaluddin, Vaishali G. Genetic divergence and principal component analysis for yield and yield components of pigeon pea [*Cajanus cajan* (L.) Millsp.] germplasm. Plant Science Today (Early Access). https:/doi.org/10.14719/pst.9079

Abstract

Pigeon pea is an important leguminous crop widely cultivated in tropical and subtropical regions for its protein-rich seeds and soil-enriching properties. The present investigation was conducted in an augmented block design during *Kharif* 2020 at the Research Farm of Banda University of Agriculture and Technology, Banda, Uttar Pradesh, India. The experimental material consisted of 73 genotypes of pigeon pea, including three checks, NDA-1, Bahar and Pant Arhar-291 to examine genetic diversity among these genotypes. Analysis of variance revealed significant genotypic variation for most of the traits among the genotypes, suggesting that the variation may be exploited through selection for the development of a variety with high-yield genetic potential. Cluster analysis grouped the genotypes into five distinct, non-overlapping clusters. Cluster I contained the highest number of genotypes (25), followed by Cluster III (19), Cluster II (14), Cluster V (9) and Cluster IV (6). Clusters II and IV had the greatest inter-cluster distance, which indicated that crossing genotypes between these clusters would result in high-quality recombinants. Principal component analysis (PCA) extracted five components that contributed 67.89 % variance to the total variation among the tested genotypes. The genotypes which showed the highest genetic diversity can be used as a parent in a hybridization program. Early flowering was observed with the genotypes IPA-10W-8-1 (74.09 days), IPA-12W-16 (78.09 days) and genotypes IPA-12W-3 (202.37 g), IPAB-10-13 (199.70 g), IPA-17W-218 (195.72 g), IPA-12W-38 (192.80 g), Local-4 (191.83 g), IPA-10W-5-8 (190.83 g) had the highest seed yield per plant. These germplasm lines can be used as potential parents for the development of high-yielding and early maturing cultivars.

Keywords: cluster analysis; pigeon pea; principal components; shelling %

Introduction

Pigeon pea [*Cajanus cajan* (L.) Millsp.], a member of the Leguminosae (Fabaceae) family, experiences 25-30 % cross-pollination (1). It is one of the main protein-rich legume crops that is cultivated extensively in tropical and subtropical locations and has a significant role in the Indian economy (2). In India, during 2020-21 pigeon pea covered an area of 6.5 M ha and produced around 5.45 metric tons of seeds (3). India is considered the primary center of origin of pigeon pea. Because it is a reliable and drought-tolerant crop, pigeon pea ensures sustained yields from marginal soils even with low inputs; as a result, it is more suited for subsistence agriculture (4).

The size of pigeon pea genome is 833.1 Mb arranged into 11 linkage groups (2). It is a multipurpose crop that is mostly grown for its edible grains. The seeds contain about 20-24 % of protein which includes high levels of important dietary amino acids including leucine (16.48 g/kg), tyrosine (14.77 g/kg) and arginine (13.51 g/kg) (5).

Through the evaluation of genetic diversity, diverse genotypes may be classified and identified from the germplasm. Genetic diversity may be defined as the degree to which heritable materials vary within a set of plants because of evolutionary influences like domestication and plant breeding (6). Geographical distances or genetic obstacles to crossing ability may cause genetic divergence among genotypes (7). The identified genetically diverse parents may be used as parents in heterosis breeding programs that exhibit a wide range of variability, helping in the isolation of transgressive segregates in advanced generations (8, 9). These genetic resources can be utilized globally in pigeon pea breeding and enhancement programs (10). While cluster analysis produces a diversity pattern, principal component analysis (PCA) is utilized to corroborate it. Principle component analysis identifies the multivariate representative sample size, analyze how principal components responded to the number of plants sampled per experimental plot and develop an approach for estimating sample size for principal components as a function of accuracy

level (11). Therefore, it was intended for the current study to classify genotypes using PCA and hierarchical cluster analysis to assess diversity (12).

To assess genetic diversity among various genotypes, genotypic divergence (D^2) analysis is helpful (13), which was later explained in detail by Rao (14). It is the approach that is most frequently used for assessing the genetic diversity of various crops. The primary methods of improvement for increased yield with resistance to severe disease, big seed size and shortened maturity have continued to be conventional breeding techniques, including single plant selection from regionally suited cultivars and pedigree selection from inter-varietal crosses (15). Understanding and knowledge of genetic variability, knowledge of genetic background information, collection and use of plant genetic resources and categorization are crucial and the foundation of crop improvement initiatives (16-18). Parental selection and the identification of genetic variability are aided by PCA and cluster analysis (19). To find genetically diverse parents for the next breeding plan, an attempt has been undertaken in the current study to evaluate the existence of genetic variation in 73 genotypes of pigeon pea using D^2 and PCA (20).

Materials and Methods

Plant materials

The experimental material comprised of 73 pigeon pea genotypes, including local landraces and three check varieties: NDA-1, Bahar and Pant Arhar-291. The elite lines were obtained from the ICAR-Indian Institute of Pulses Research (IIPR), Kanpur, a center with an advanced pigeon pea breeding program. Local landraces were collected from the Banda region to assess their adaptability to regional agro-climatic conditions and potential farmer-preferred traits. A complete list of the genotypes used in the study is presented in Table 1.

Study research site

The field experiment was conducted during the Kharif season of 2020 at the Research Farm of Banda, University of Agriculture and Technology, Banda, India. The precise location of the research farm is identified by its geographic coordinates: Latitude 25.533743°N, Longitude 80.338042°E and it is situated at an altitude of 113 meters above mean sea level. The soil of the experimental field is clay loam and the pH is 7.4-7.7.

The maximum temperature ranged from 17.30 °C to 39.0 °C while the minimum temperature was from 10.41 °C to 27.58 °C, the maximum rainfall was received up to 158 mm and the average relative humidity varied from 16.90 % to 86.67 % during the crop period.

Table 1. List of pigeon pea genotypes including check used in the present study

SN.	Genotype	Source	S.N.	Genotype	Source
1	IPA-12W-51	IIPR, Kanpur	38	IPAB-11-20-2	IIPR, Kanpur
2	IPA-12-100	IIPR, Kanpur	39	LOCAL-5	Patwan, Banda
3	IPA-12W-81	IIPR, Kanpur	40	IPAB-11-9-1	IIPR, Kanpur
4	IPDA-4	IIPR, Kanpur	41	IPA-12W-45	IIPR, Kanpur
5	LOCAL-1	Baberu, Banda	42	IPAB-10-13	IIPR, Kanpur
6	IPAB-11-1-3	IIPR, Kanpur	43	DBN-711	IIPR, Kanpur
7	IPM-16-1	IIPR, Kanpur	44	IPA-12W-3	IIPR, Kanpur
8	IPA-10W-5-8	IIPR, Kanpur	45	IPAB-11-19-3	IIPR, Kanpur
9	IPAB-11-10-1	IIPR, Kanpur	46	IPAB-10-37	IIPR, Kanpur
10	IPAB-11-8-1	IIPR, Kanpur	47	IPAB-11-1-8	IIPR, Kanpur
11	IPA-12W-49	IIPR, Kanpur	48	IPA-12W-385	IIPR, Kanpur
12	IPAB-11-16-1	IIPR, Kanpur	49	UPAS-120	IIPR, Kanpur
13	LOCAL-2	Chitrakute, Banda	50	IPA-12W-82	IIPR, Kanpur
14	IPA-12W-220	IIPR, Kanpur	51	IPAB-81-9	IIPR, Kanpur
15	LOCAL-3	Marka, Banda	52	IPAB-10-10	IIPR, Kanpur
16	IPA-12W-227	IIPR, Kanpur	53	IPAB-11-14-1	IIPR, Kanpur
17	IPAB-11-1-4	IIPR, Kanpur	54	IPA-12W-16	IIPR, Kanpur
18	IPAB-81-5	IIPR, Kanpur	55	IPAM-15-2	IIPR, Kanpur
19	IPA-12W-382	IIPR, Kanpur	56	IPA-10W-20-3	IIPR, Kanpur
20	IPA-12W-43	IIPR, Kanpur	57	IPA-12-37	IIPR, Kanpur
21	IPA-12W-84	IIPR, Kanpur	58	IPAB-16-17-1	IIPR, Kanpur
22	PKV-TARS	IIPR, Kanpur	59	IPA-12W-50	IIPR, Kanpur
23	IPA-12W-208	IIPR, Kanpur	60	IPA-12W-398	IIPR, Kanpur
24	IPA-12W-53	IIPR, Kanpur	61	IPA-12W-45	IIPR, Kanpur
25	LOCAL-4	Majganva, Banda	62	IPAB-11-13-1	IIPR, Kanpur
26	IPAB-11-6-1	IIPR, Kanpur	63	IPA-10W-20-3	IIPR, Kanpur
27	IPAB-11-20-1	IIPR, Kanpur	64	IPA-17B-12	IIPR, Kanpur
28	IPA-10W-2-7	IIPR, Kanpur	65	IPAB-81-6	IIPR, Kanpur
29	IPAB-11-9-1	IIPR, Kanpur	66	LOCAL-6	Tindwari, Banda
30	IPAB-1-6	IIPR, Kanpur	67	IPA-10W-8-1	IIPR, Kanpur
31	IPAM-16-2	IIPR, Kanpur	68	IJT-50	IIPR, Kanpur
32	IPA-12W-38	IIPR, Kanpur	69	IPA-12W-47	IIPR, Kanpur
33	IPAB-1-3	IIPR, Kanpur	70	IPA-12W-389	IIPR, Kanpur
34	IPA-12W-46	IIPR, Kanpur	71	NDA-1 (Check-1)	IIPR, Kanpur
35	IPAB-10-9	IIPR, Kanpur	72	Bahar (Check-2)	IIPR, Kanpur
36	IPA-12W-395	IIPR, Kanpur	73	Pant Arhar-291 (Check-3)	IIPR, Kanpur
37	IPA-17W-218	IIPR, Kanpur	-	-	-

Experimental design and observations recorded

An augmented block design was used to set up the experiment. Each genotype was sown in a net plot area of 6 m^2 with a 4 m row length and spacing of 30 cm between the rows. The data were collected on five randomly selected plants from each genotype for plant height (PH), number of primary branches (NPB), number of secondary branches (NSB), number of pods per plant (NPP), number of seeds per pod (NSP), pod length (PL), 100-seed weight (SW), biological yield per plant (BYP), harvest index (HI) and seed yield per plant (SY). Days to 50 % flowering (DF) and days to maturity (DM) were recorded on a plot basis. Shelling Percentage (%) was calculated by the formula Shelling (%) = Grain weight (g)/Pod weight (g) × 100. Harvest index was calculated through the formula Harvest index (%) = Seed yield per plant (g)/Biological yield per plant (g) × 100.

Statistical analysis

The statistical analysis was done using R software (version 4.5.0). The 'ggplot2' and 'factoextra' packages were used for PCA analysis. Hierarchical clustering and dendrograms were generated with the 'hclust' function to identify genetic diversity in the studied materials.

Results and Discussion

Analysis of variance

Wide genetic variation and efficient selection strategies are the keys to the success of any breeding programs that make it possible to exploit existing genetic resources (21-23). The analysis of variance revealed highly significant differences among the genotypes with respect to almost all the characters under study (Table 2). This indicated that there is significant inherent genetic variability among the genotypes under study, which provides ample scope for identifying genotypes with desirable characters to improve yield, provided the material is subjected to sensible selection pressure similar findings were also reported in the previous studies (24-26). Significant variations between treatments were also observed for the number of days to germination, days to flower initiation, days to 50 % flowering, days to 100 % flowering, days to first pod appearance, days to maturity, plant height, height of first pod and grain production (27).

Distribution of genotypes into clusters

Mahalanobis' D^2 statistics were computed between all possible pairs of 73 pigeon pea genotypes and the genetic diversity present among the genotypes was assessed. In cluster analysis, 73 germplasm lines were grouped into five different non-overlapping clusters in which cluster I had the maximum number of genotypes (25) followed by cluster III (19 genotypes), cluster II (14 genotypes), cluster V (09 genotypes) and cluster IV (06 genotypes) (Table 3). Twenty-three genotypes of pigeon pea were grouped into five clusters by the Dendrogram method (Fig. 1). Rao et al. (28) grouped 100 germplasm lines of rice in 4 clusters, in which cluster 1 had the minimum while cluster 4 had the maximum number of individuals. The individuals gathered in a particular cluster were genetically similar to one another within the cluster (28). In the present study, the genotypes were classified into five different clusters, providing an ample opportunity to identify diverse parental lines for the crossing program in pigeon pea. It is essential to introduce new genetic variations or integrate genes from genetically diverse parent plants to utilize observed genetic diversity found among the examined pigeon pea genotypes (21). Hybridization between lines selected from different clusters is likely to produce more heterotic hybrids.

Table 2. Analysis of variance of augmented block design for 13 characters in pigeon pea genotypes

Characters	Sources of variation					
	Blocks	Treatment	Genotype	Checks	Genotype *vs.* check	Error
d.f.	6	72	69	2	1	12
DF	34.19	1027.66*	816.64*	8511.04*	621.90*	21.04
DM	94.15	4440.66*	3221.71*	48560.33*	309.36*	35.11
PH	46.54*	170.20*	166.60*	133.16*	492.72*	14.22
NPB	4.61*	6.11*	5.99*	6.41*	14.15*	1.5
NSB	15.76*	22.46*	19.31*	110.93*	62.87*	62.87
NPP	6964.17*	7523.04*	7110.45*	1627.72	47782.45*	898.83
NSP	0.031*	0.013	0.012	0.047*	0.047	0.01
PL	0.43*	0.19	0.17	0.95*	0.95	0.95
100-SW	1.49	3.04*	2.77*	13.24*	0.93	0.84
BYP	311.41*	1717.85*	1732.07*	544.16*	3084.24*	80.11
HI	2.73	22.63*	21.35*	62.49*	31.87*	3.77
Shelling %	69.52*	41.83*	32.94*	327.97*	83.02*	10.27
SY	83.37	563.24*	538.48*	1693.13*	11.87	31.55

*Indicates significant at 5 %; DF: days to 50 % flowering; DM: days to maturity; PH: plant height; NPB: number of primary branches; NSB: number of secondary branches; NPP: number of pods per plant; NSP: number of seeds per pod; PL: pod length; SW: 100-Seed weight; BYP: biological yield per plant; HI: harvest index; SY: seed yield per plant.

Table 3. Clustering pattern of 73 pigeon pea genotypes (including three checks) for yield and its attributing traits

Cluster	Number of genotypes	Genotypes
I	25	NDA-1, IPAB-11-1-3, LOCAL-3, IPA-12W-81, IPAB-11-16-1, IPAB-11-1-8, IPAB-10-13, IPA-12W-3, IPA-10W-20-3, UPAS-120, IPAB-11-13-1, IPA-12W-277, IPA-12W-208, Pant Arhar-291, IPA-12W-385, IPAB-11-20-2, IPAB-11-1-4, IPAM-16-1, IPA-17W-218, IPAB-11-6-1, IPAM-16-2, IPAB-11-14-1, IPAB-10-37, IPA-12W-385, IJT-50
II	14	LOCAL-2, LOCAL-4, IPA-12W-389, IPA-12W-82, IPA-12W-53, LOCAL-6, IPA-12-37, IPA-12W-220, Bahar, IPA-12W-38, IPA-12W-47, IPA-12W-50, IPA-12W-49, IPAB-81-5
III	19	IPA-10W-8-1, IPAB-11-10-1, IPA-12W-46, PKV- Tars, IPAB-11-9-1, DBN-711, IPAM -15-2, IPAB- 4, IPA-12W-45, IPA-10W-5-8, IPAB-10-9, IPAB-11-20-1, IPAB-11-9-1, IPA-12W-51, IPAB-11-19-3, IPAB-81-9, IPA-10W-20-3, IPA-12W-43, IPAB-81-6
IV	6	IPAB-11-8-1, IPA-12W-16, IPA-12W-100, IPA-12W-84, IPAB-10-10, IPA-17B-12,
V	9	LOCAL-1, IPA-12W-398, IPA-12W-395, IPA-12W-45, IPA-10W-2-7, LOCAL-5, IPAB-16-17-1, IPAB-1-6, IPAB-1-3

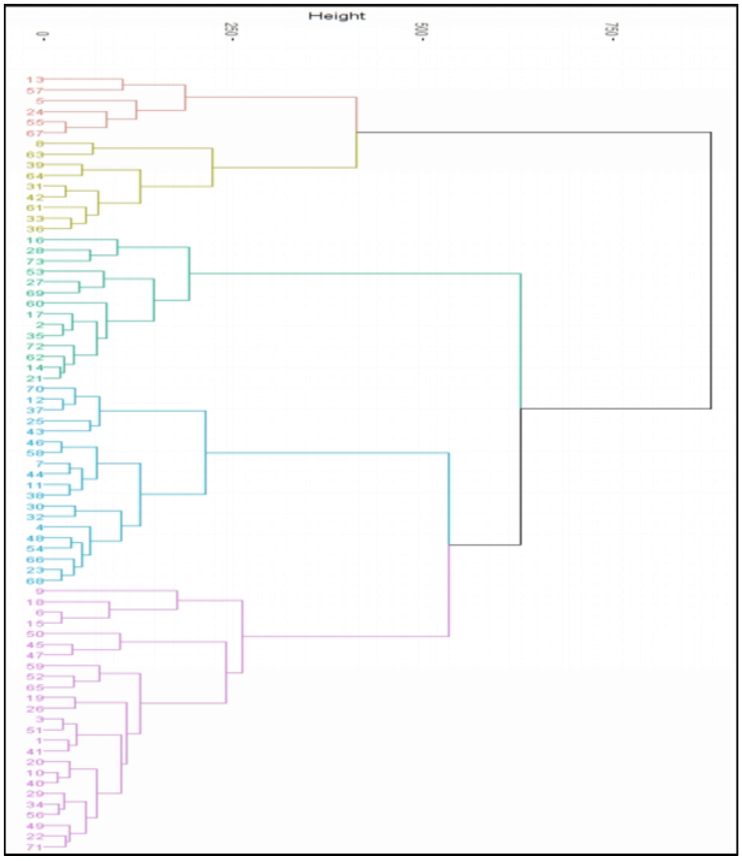

Fig. 1. Dendrogram showing the relationship between 73 genotypes including three checks of pigeon pea genotypes. 1- NDA-1, 2- BAHAR, 3- PANT ARHAR-291, 4- IPA-12W-51, 5- IPA-12W-100, 6- IPA-12W-81, 7- IPAB-4, 8- LOCAL-1, 9- IPAB-11-1-3, 10- IPAM-16-1, 11- IPA-10W-5-8, 12- IPAB -11-10-1, 13- IPAB-11-8-1, 14- IPA-12W-49, 15- IPAB-11-16-1, 16- LOCAL-2, 17- IPA-12W-220, 18- LOCAL-3, 19- IPA-12W-277, 20- IPAB-11-1-4, 21- IPAB-81-5, 22- IPA-12W-382, 23- IPA-12W-43, 24- IPA-12W-84, 25- PKV-TARS, 26- IPA-12W-208, 27- IPA-12W-53, 28- LOCAL-4, 29- IPAB-11-6-1, 30- IPAB-11-20-1, 31- IPA-10W-2-7, 32- IPAB-11-9-1, 33- IPAB-1-6, 34- IPAM-16-2, 35- IPA-12W-38, 36- IPAB-1-3, 37- IPA-12W-46, 38- IPAB-10-9, 39- IPA -12W-395, 40- IPA-17W-218, 41- IPAB-11-20-2, 42- LOCAL-5, 43- IPAB-11-9-1, 44- IPA-12W-45, 45- IPAB-10-13, 46- DBN-711, 47- IPA-12W-3, 48- IPAB-11-19-3, 49- IPAB-10-37, 50- IPAB-11-1-8, 51- IPA-12W-385, 52- UPAS-120, 53- IPA-12W-82, 54- IPAB-81-9, 55- IPAB-10-10, 56- IPAB-11-14-1, 57- IPA-12W-16, 58- IPAM-15-2, 59- IPA-10W-20-3, 60- IPA-12W-37, 61- IPAB-16-17-1, 62- IPA-12W-50, 63- IPA-12W-398, 64- IPA-12W-45, 65- IPAB-11-13-1 , 66- IPA-10W-20-3, 67-IPA-17B-12, 68- IPAB-81-6, 69- LOCAL-6, 70- IPA-10W-8-1, 71- IJT-50, 72- IPA-12W-47, 73- IPA-12W-389.

Inter-cluster and intra-cluster distance

The intra and inter-cluster distances between all possible pairs of five clusters were computed and presented in Table 4. While the minimum inter-cluster distance demonstrated a close link between the groupings, the largest inter-cluster distance indicated broad diversity (29). The maximum intra-cluster distance was found in cluster IV (111.35) followed by cluster V (93.57), cluster I (89.82), cluster II (79.34) and minimum intra-cluster distance was found in cluster III (72.72). The maximum inter-cluster distance was found from cluster II to IV (280.14), followed by cluster I to IV (276.68), cluster I to V (243.34), cluster IV to V (181.72), cluster III to V (176.20), cluster III to IV (176.14), cluster II to III (173.01), cluster I to II (161.36) and minimum inter -cluster distance was found between I and III (135.01). Similar results were also obtained in the previous studies (30, 31) in which it is reported that a high magnitude of inter-cluster distance as compared to intra-cluster distance. The presence of substantial inter-cluster distance as compared to intra-cluster distance indicated that a considerable amount of genetic diversity exists among the examined genotypes (32). To increase the probability of obtaining desirable segregants, preference is given to optimal genetic divergence among the parental lines for hybridization. Therefore, it is suggested that the genotypes from clusters with maximum inter-cluster distance may be crossed to get superior segregants in succeeding generations.

Table 4. The average intra and inter-cluster distances for five clusters in pigeon pea germplasm

Cluster	I	II	III	IV	V
I	**89.82**	161.36	135.01	276.68	243.34
II		**79.34**	173.01	280.11	171.58
III			**72.72**	176.14	176.2
IV				**111.35**	181.72
V					**93.57**

Bold figures indicate the intra-cluster distance into different genotypes is divided into five clusters.

Cluster means for different characters

The cluster means for all 13 characters are presented in Table 5 and it showed that among all clusters, cluster I rendered the highest cluster mean values for the characters viz. pod length (6.59), number of primary branches per plant (25.00), harvest index (43.93) and number of pods per plant (835.41). However, the highest cluster means for the number of seed per plant (3.60), 100 seed weight (10.15) and shelling % (69.02) reported in cluster III; for days to 50 % flowering (157.16), seed yield per plant (178.28) and biological yield per plant (413.90) in cluster II; for days to maturity (272.01) and plant height (228.52) in cluster V; for number of secondary branches (66.56) in cluster IV. The maximum yield was observed for the genotypes of cluster II and a minimum of cluster IV. It is clearly indicated from the result that neither of any clusters has genotypes with all desirable traits. Therefore, crossing between genotypes of

Table 5. Cluster mean for different characters in pigeon pea germplasm

Cluster	I	II	III	IV	V
DF	101.11	157.16**	95.98*	103.76	152.2
DM	156.85	271	147.94*	156.26	272.01**
PH	222.49	222.62	222.37	222.29*	228.52**
NPB	25.00**	24.56	24.35	22.50*	24.58
NSB	63.25*	65.34	65.23	66.56**	65.15
NPP	835.41**	789.93	721.49	586.73*	646.33
NSP	3.56	3.55*	3.60**	3.56	3.56
PL	6.59**	6.3	6.19	6.09*	6.17
100-SW	9.94	9.75	10.15**	8.64*	9.75
BYP	382.49	413.90**	376.82	307.44*	363.02
HI	43.93**	43.33	43.93	40.92*	43.25
Shelling %	68.21	66.45	69.02**	65.54*	68.62
SY	168.02	178.28**	165.46	123.54*	157.34

*Lowest value, **Highest value.

different clusters is desirable to develop superior genotypes. Similar observations on the utility of cluster-based hybridization were reported (33, 34).

This result indicates the importance of the selection of genotypes from the corresponding clusters in hybridization programs for effecting improvement in the respective traits (34). The most desirable pigeon pea genotypes were identified for all 13 characters in the 73 genotypes through cluster analysis (Table 6).

Principal component analysis

This method was first developed by Pearson and later used by Hotteling. Now it is being currently used by several workers to select superior genotypes (35, 36). PCA is a powerful tool in modern data analysis because this is a well-known multivariate statistical technique that is used to identify the minimum number of components, which may rank genotypes based on PC scores and can explain the most variability out of all variability (34, 37).

Principal components are generally estimated either from a correlation matrix or a covariance matrix. Considering the importance of PCA, this study is conducted on pigeon pea genotypes to the identification of the yield and yield-related traits responsible for the yield differences among the pigeon pea genotypes. The estimates on PCA for 13 quantitative traits are presented in Table 6. The perusal of the table revealed that the recorded data from 13 traits were transformed into 13 principal components. The first five principal components with an eigenvalue of more than one showed 67.89 % of the total variation present in studied genotypes. PC1 alone contributed 20.22 % of the variation, followed by PC2 (17.04 %), PC3 (12.00 %), PC4 (10.66 %) and PC5 (7.97 %). The eigenvalues of the first five components were 2.628, 2.215, 1.560, 1.386 and 1.034, respectively. These results suggest that a substantial proportion of trait variation can be explained by a small number of components. Similar trends have been reported in pigeon pea (38, 39). The key traits contributing to variation in each principal component are providing insights into which phenotypic variables most influence genetic diversity in the genotypes studied Table 6.

Principal Component I (PC1), which accounted for the highest proportion of variability among genotypes, showed high loadings for seed yield per plant, number of pods per plant, 100-seed weight, biological yield per plant and number of primary branches per plant. Days to 50 % flowering, days to maturity, pod length and biological yield per plant were mainly contributed to PC II while pod length, harvest index (%) number of seeds per pod and shelling (%) to PC III (Table 7). These results indicated that the above traits contributed the most to the divergence and carried most of its variability. The yield level was thus decided by these five components, which was the weighted average of the characters. Similar findings regarding the influence of secondary branches, pods per plant and plant height were reported (31).

The number of principal components to retain is typically determined using the Kaiser criterion, which considers components with eigenvalues greater than one as significant. In this study, only the first five PCs met this threshold, collectively accounting for most of the phenotypic variation among pigeon pea genotypes for yield-related traits. Therefore, traits associated with these five components should be prioritized in pigeon pea improvement programs (38, 40). The principal component biplot (Fig. 2) graphically represents all 13 traits as vectors, illustrating their relative contributions to genotype differentiation and helping to identify promising trait combinations for selection.

Principal component analysis (PCA) is an effective multivariate technique to find and determine the independent main components that govern plant attributes individually. PCA bioplot measures the association between the traits to choose effective characters using an indirect selection of superior genotypes (41, 42). The gap within traits indicates the association among them, like if two variables, away from the origin, having an acute angle (less than 90°) were positively correlated, as biological yield PC1 and PC2 indicated the contribution of these traits toward divergence (Fig. 3). The length of the vector represented the magnitude of that character. The scree plot showed that the major contribution towards divergence was due to PC1 followed by PC2, PC3, PC4 and PC5 (Fig. 3). A scatter plot of 13 characters using PC1 and PC2 is shown in (Fig. 4).

Table 6. The most desirable pigeon pea genotypes were identified for 13 characters

Characters	Genotypes
DF	IPA-10W-8-1, IPA-12W-16
DM	IPA-10W-8-1, IPA-12W-16, IPA-10W-5-8, IPAB-11-1-4, IPAB-11-20-1
PH	IPA-10W-20-3, IPA-12-37, IPA-12W-277, IPAB-81-5, IPA-12W-43, IPAB-11-9-1
NPB	IPAB-10-37, IPA-12W-3, IPAB-11-20-2, IPA-12W-46, IPAB-11-20-2, Local-5
NSB	IPA-12W-208, Local-4, IPA-12W-53, IPA-10W-2-7, IPAB-11-20-1, IPA-12W-38
NPP	IPA-12W-3, IPAB-10-13, IPAB-11-1-8, UPAS-120, IPAB-10-37, IPAM-16-1
NSP	IPAM-15-2, IPAB-11-19-3, IPAB-1-6, IPA-12W-385, IPA-12W-84, IPAB-11-20-1
PL	UPAS-120, IPA-12W-82, Local-1, Local-3, IPAB-11-6-1
100-SW	IPA-12W-395, IPA-12W-81, IPA-17W-218, IPA-12W-45, IPA-12W-3, IPA-12-37
BYP	Local-2, Local-4, IPAB-11-1-4, IPA-12W-38
HI	IPAB-10-13, IPA-12W-3, PKV-Tars, IPA-12W-47, IPA-12W-277
Shelling %	IPAB-11-19-3, IPAB-11-1-8, IPAB-10-37
SY	IPA-12W-3, IPAB-10-13, IPA-17W-218, IPA-12W-38, Local-4, IPA-10W-5-8

Table 7. Principal component analysis for 13 quantitative traits in pigeon pea germplasm

Characters	PC1	PC2	PC3	PC4	PC5
Eigenvalue	2.628	2.215	1.560	1.386	1.034
Variance (%)	20.222	17.045	12.004	10.662	7.957
Cumulative Variance (%)	20.222	37.268	49.273	59.935	67.893
Eigenvectors					
DF	0.014	0.603	-0.238	-0.221	0.079
DM	-0.004	0.604	-0.225	-0.231	0.102
PH	-0.009	0.116	0.195	-0.254	-0.733
NPB	-0.333	-0.108	-0.011	-0.247	-0.066
NSB	-0.021	0.078	0.188	-0.207	-0.323
NPP	-0.388	0.056	-0.243	0.319	-0.240
NSP	-0.194	-0.148	0.019	-0.455	0.168
PL	-0.175	-0.235	-0.518	-0.199	-0.310
100-SW	-0.361	-0.078	0.129	-0.121	0.339
BYP	-0.357	0.210	-0.237	0.413	-0.056
HI	-0.340	0.056	0.477	-0.195	0.065
Shelling %	-0.109	-0.247	-0.394	0.357	0.173
SY	-0.531	0.197	0.181	0.166	0.015

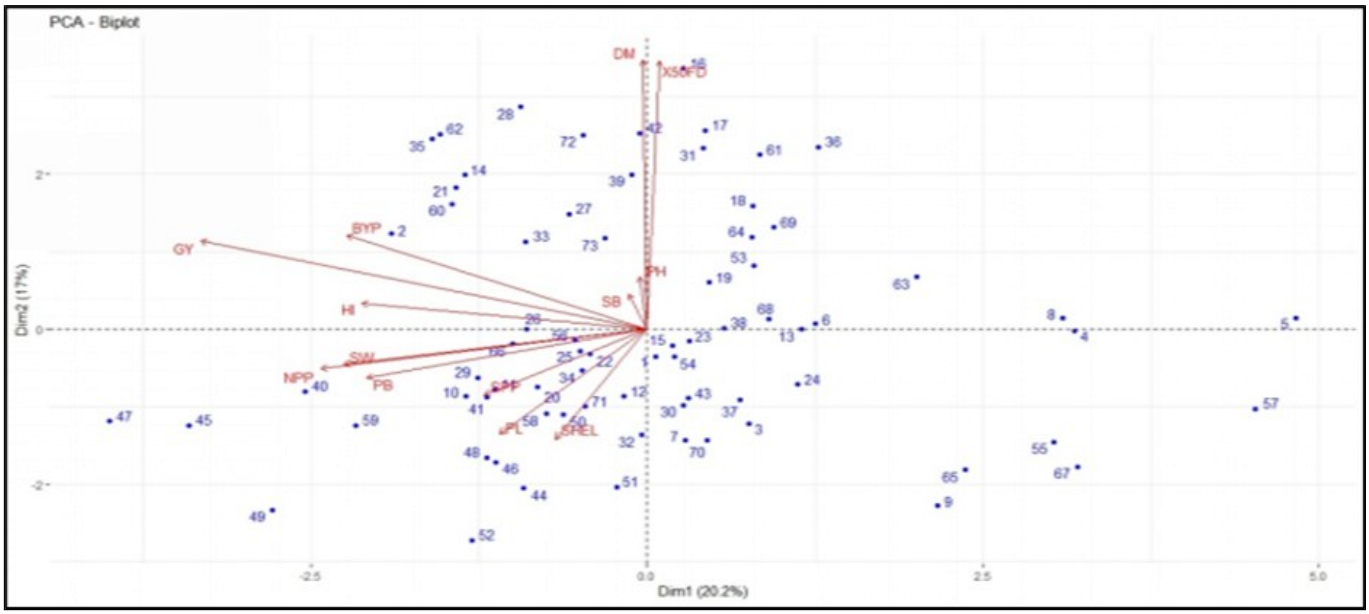

Fig. 2. PCA Biplot of 73 genotypes, including three checks of pigeon pea genotypes for yield and yield-related traits.

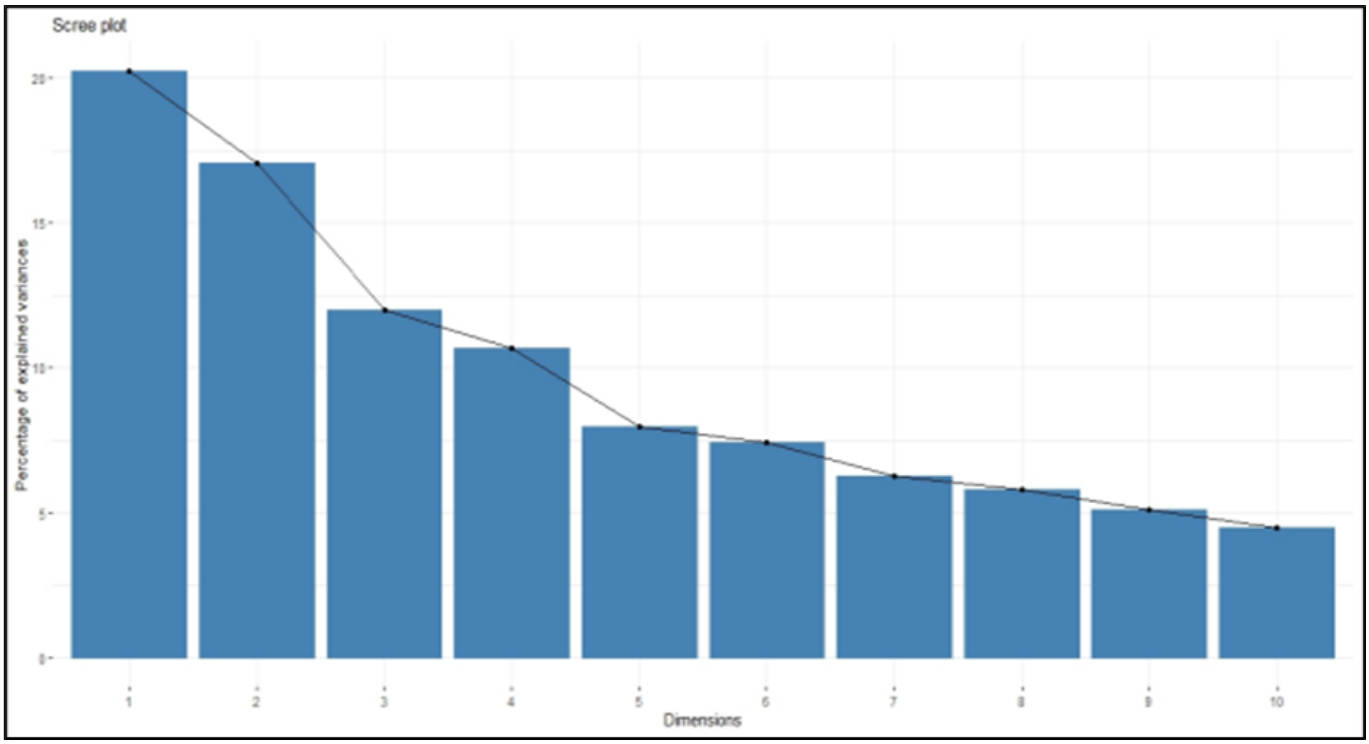

Fig. 3. Screen plot reflected the eigen value of different principal component and the percent cumulative variability as shown by PCA of pigeon pea germplasm based on quantitative traits.

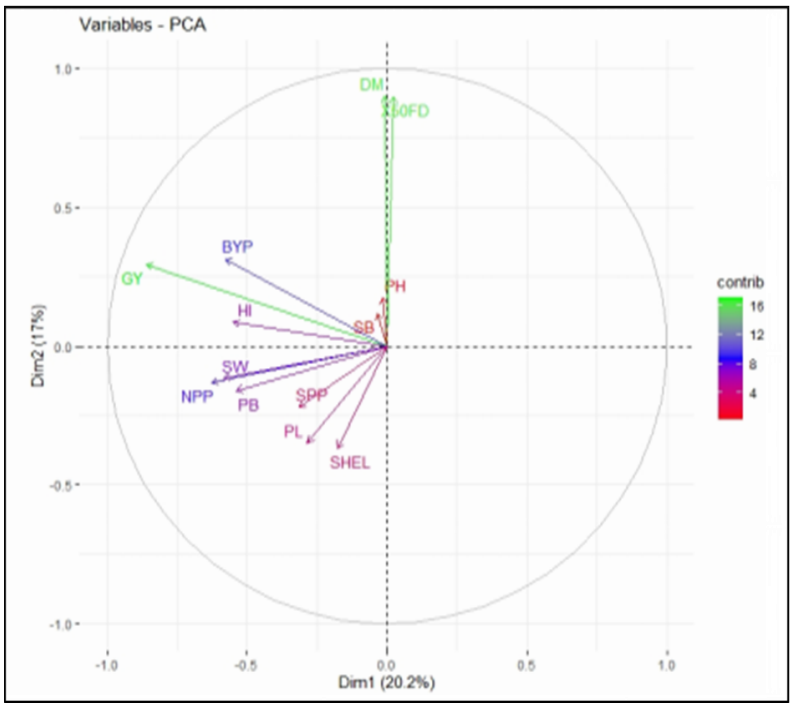

Fig. 4. Scatter plot of 13 quantitative characters of pigeon pea using PC1 and PC2.

Conclusion

Mahalanobis D^2 analysis revealed genetic divergence among 73 pigeon pea genotypes and grouped into five clusters by the dendrogram method. Based on divergence classes, clusters II and IV showed the highest amount of diversity. This suggested that crossing among genotypes from different clusters may result in a wide array of variability for exercising effective selection and could increase the probability to recover desirable segregants in subsequent generations. The genotypes with desirable attributes from the corresponding cluster could be further evaluated for isolating high-yielding and early maturing genotypes employing mass selection and pedigree selection. PCA showed that five PCs were found significant and contributed 67.893 % of the total cumulative variability among the different genotypes. Based on these components, the genotypes are selected which would be utilized for the further breeding program for developing high-yielding cultivars.

Acknowledgements

The authors are grateful to the Department of Genetics & Plant Breeding, BUAT, Banda and Pulses Improvement Project CEDA, Govt. of U.P., India for providing financial assistance to carry out this study. The authors are also thankful to Dr. Mukul Kumar for providing infrastructural facilities and IIPR-Kanpur for the supply of seed during the research work.

Authors' contributions

AK contributed to the conceptualization and methodology, performed data collection and curation, prepared the initial draft of the manuscript. NK conducted statistical analysis, prepared the manuscript. VS conceptualized the problem and methodology, performed statistical analysis. HSN reviewed and edited the manuscript, provided resources to carry out the experiments. SKS reviewed and edited the manuscript. K contributed to the conceptualization and methodology, provided resources to carry out the experiments. VG reviewed and edited the manuscript, provided resources to carry out the experiments. All authors read and approved the final manuscript.

Compliance with ethical standards

Conflict of interest: Authors do not have any conflict of interests to declare.

Ethical issues: None

References

1. Saxena K, Bohra A, Choudhary AK, Sultana R, Sharma M, Pazhamala LT, et al. The alternative breeding approaches for improving yield gains and stress response in pigeon pea (*Cajanus cajan*). Plant Breed. 2021;140(1):74-86. https://doi.org/10.1111/pbr.12863
2. Varshney RK, Chen W, Li Y, Bharti AK, Saxena RK, Schlueter JA, et al. Draft genome sequence of pigeon pea (*Cajanus cajan*), an orphan legume crop of resource-poor farmers. Nat Biotechnol. 2012;30(1):83-9. https://doi.org/10.1038/nbt.2022
3. FAO Corporate Statistical Database (FAOSTAT); 2022. https://www.fao.org/faostat/en/#data
4. Géofroy K, Gustave D, Laura EYL, Relique IA, Saxena RK, Varshney RK, et al. Agro-morphological characterization of pigeon pea (*Cajanus cajan* L. Millspaugh) landraces grown in Benin: implications for breeding and conservation. J Plant Breed Crop Sci. 2020;12(1):34-49. https://doi.org/10.5897/JPBCS2019.0836
5. Yohane EN, Shimelis H, Laing M, Mathew I, Shayanowako A. Genotype-by-environment interaction and stability analyses of grain yield in pigeon pea (*Cajanus cajan* L. Millspaugh). Acta Agri Scand B Soil Plant Sci. 2021;71(3):145-55. https://doi.org/10.1080/09064710.2020.1859608
6. Hossain MM, Azad MAK, Eaton TEJ, Siddquie MNA. Assessment of different traits to evaluate genetic divergence in some wheat (*Triticum aestivum* L.) genotypes under late sowing conditions. Am J Plant Sci. 2021;12(5):734-47. https://doi.org/10.4236/ajps.2021.125050
7. Singh NK, Mishra SN. Genetic divergence in oat germplasm. Indian

J Genet. 1996;56(3):272-7.

8. Sujit MK, Belsariya N, Pandey VK. Studies on genetic diversity for yield and yield attributing traits in pigeon pea (*Cajanus cajan* L. Millsp) germplasm. Pharm Innov. 2023;12(3):5272-7.
9. Upadhyaya HD. Germplasm management for enhanced genetic gains. IJPGR. 2016;29(3):261-4. https://doi.org/10.5958/0976-1926.2016.00041.3
10. de Souza, Cargnelutti Filho A, Toebe M, Bittencourt KC. Sample size and genetic divergence: a principal component analysis for soybean traits. Eur J Agron. 2023;149:126903. https://doi.org/10.1016/j.eja.2023.126903
11. Hemavathy AT, Bapu JR, Priyadharshini C. Principal component analysis in pigeon pea (*Cajanus cajan* L. millsp.). Electron J Plant Breed. 2017;8(4):1133-9. https://doi.org/10.5958/0975-928X.2017.00165.X
12. Mahalanobis PC. On the generalized distance in statistics. Proc Nat Ins India. 1936;2:49-55.
13. Rao CR. Advanced statistical methods in biometrical research. 1st edition. New York: John Wiley and Sons; 1952. p. 236-72.
14. Saxena KB, Sawargaonkar SL. First information on heterotic groups in pigeon pea (*Cajanus cajan* L. Millsp.). Euphytica. 2014;200:187-96. https://doi.org/10.1007/s10681-014-1142-0
15. Khan SA, Iqbal J, Khurshid H, Saleem N, Rabbani MA, Zia M, et al. The extent of intra-specific genetic divergence in the *Brassica napus* L. population was estimated through various agro morphological traits. Eur J Acad Res. 2014;2:2255-75. https://doi.org/10.14419/ijbas.v3i2.2336
16. Syafii M, Cartika I, Ruswandi D. Multivariate analysis of genetic diversity among some maize genotypes under maize-albizia cropping system in Indonesia. Asian J Crop Sci. 2015;7:244-55. https://doi.org/10.3923/ajcs.2015.244.255
17. Mofokeng MA, Bello Z, Mashingaidze K, Gerrano AS. Morphological plasticity of economical traits in pigeon pea genotypes grown in South Africa. Intl J Agric Biol. 2022;28(2):75-84.
18. Mohammed A, Tesso B, Ojiewo C, Ahmed S. Assessment of genetic variability and heritability of agronomic traits in Ethiopian chickpea (*Cicer arietinum* L) landraces. BSJ Agri. 2019;2:10-5.
19. Zavinon F, Adoukonou-Sagbadja H, Bossikponnon A, Dossa H, Ahanhanzo C. Phenotypic diversity for agro-morphological traits in pigeon pea landraces (*Cajanus cajan* L. Millsp.) cultivated in southern Benin. Open Ag. 2019;4(1):487-99. https://doi.org/10.1515/opag-2019-0046
20. Chauhan N, Singh I, Kumar H, Vij S. Evaluation of genetic diversity of advanced pigeonpea (*Cajanus cajan* L. Millspaugh) interspecific derivatives using Mahalanobis D^2 and selection of elite lines using multi-trait genotype-ideotype distance index. Int J Bio-Resour Stress Manag. 2025;16(1). https://doi.org/10.23910/1.2025.5916
21. Mauriya AK, Kumar V, Verma RB, Sahu R, Hashim M. Impact of cluster frontline demonstration on increasing productivity and profitability of pigeonpea in Bihar. J Food Legumes. 2024;37(2):211-4. https://doi.org/10.59797/jfl.v37.i2.197
22. Kumar A, Kumar H, Singh CM, Kumar M, Sharma V, Kumar S, et al. Multivariate analysis for elucidating genetic diversity of chickpea (*Cicer arietinum* L.) germplasm using agro-morphological traits. Legume Res. 2025;48(4):590-6. https://doi.org/10.18805/LR-4956
23. Choudhary M, Gothwal DK, Kumawat KR, Kumawat R, Kumar O, Bajya M. Cluster and principal component analysis in fenugreek (*Trigonella foenum-graecum* L.) genotypes based on yield and yield related traits. Pharma Innov J. 2022;11(5):1723-8.
24. Alemayo GT, Gurmu GN, Singh BCS, Getachew ZE. Genetic variability of chickpea (*Cicer arietinum* L.) genotypes for yield and yield components in West Shewa, Ethiopia. Plant Cell Biotechnol Mol Biol. 2021;22(35-36):307-31.
25. Kumar K, Anjoy P, Sahu S, Durgesh K, Das A, Tribhuvan KU, et al. Single trait versus principal component-based association analysis for flowering related traits in pigeon pea. Sci Rep. 2022;12(1):10453. https://doi.org/10.1038/s41598-022-14192-7
26. Kumar A, Kumar H, Sharma V, Kamaluddin. Estimation of genetic parameters, selection indices and association analysis of seed yield and its component traits in chickpea (*Cicer arietinum* L.). Legume Res. 2021;47(3):1-9. https://doi.org/10.18805/LR-4506
27. Talekar SC, Praveena MV, Satish RG. Genetic diversity using principal component analysis and hierarchical cluster analysis in rice. Int J Plant Sci. 2022;17(2):191-6. https://doi.org/10.15740/HAS/IJPS/17.2/191-196
28. Rao PJM, Malathi S, Reddy DVV, Upender M. Genetic studies of association and path coefficient analysis of yield and its component traits in pigeon pea (*Cajanus cajan* L. Millsp.). Int J Sci Res Publ 2013;3(8):1-5.
29. Devi M, Jyothula DPB, Krishnaveni B, Rao VS. Genetic divergence studies in rice (*Oryza sativa* L.) hybrids for yield, yield component traits and quality parameters. Int J Curr Microbiol Appl Sci. 2019;8(6):1577-83. https://doi.org/10.20546/ijcmas.2019.806.189
30. Rekha R, Prasanthi L, Sekhar MR, Priya MS. Variability, character association and path analysis for yield and yield attributes in pigeon pea. Electron J Plant Breed. 2013;4(4):1331-5.
31. Gaur AK, Verma SK, Sharma RK, Bhardwaj R, Bisht C, Deep H, et al. Elucidation of genetic mechanism governing heterosis in pigeonpea (*Cajanus cajan* L. Millspaugh). Legume Res. 2025;1:6.
32. Satapathy B, Panigrahi KK. Assessment of genetic divergence in pigeon pea (*Cajanus cajan* L.). Trends Biosci. 2014;7(19):3001-5.
33. Bhatt A, Verma SK, Panwar RK, Reddy SK. Assessment of crop genetic diversity utilizing multivariate analysis in pigeonpea (*Cajanus cajan* L. Millsp.) genotypes. Legume Res. 2024;1:8. https://doi.org/10.18805/LR-5375
34. Reddy VG, Jayalakshmi V, Sreenivas T. Genetic divergence studies in pigeon pea inbreeds. Electron J Plant Breed 2015;6(2):515-20.
35. Morrison DE. Multivariate statistical methods. 2nd edition. 4th print. Tokyo: McGraw-Hill Kogakusha Ltd; 1978. p. 1-85.
36. Vianna VF, Unêda-Trevisoli SH, Desidério JA, Santiago SD, Charnai K, Ferreira Júnior JA. The multivariate approach and influence of characters in selecting superior soybean genotypes. Afr J Agric Res. 2013;8(30):4162-9.
37. Li X, Zhang C, Behrens H, Holtz F. Calculating amphibole formula from electron microprobe analysis data using a machine learning method based on principal components regression. Lithos. 2020;362:105469. https://doi.org/10.1016/j.lithos.2020.105469
38. Anderson TW. An introduction to multivariate analysis. New Delhi: Wiley Eastern Pvt. Ltd; 1972.
39. Alagu Palamuthirsolai M, Vijayabharathi A. Genetic diversity analysis of pigeon pea (*Cajanus cajan* L. Millsp.) under cold stress conditions. J Pharmacogn Phytochem. 2020;9(3):2048-53. https://doi.org/10.22271/phyto.2020.v9.i3ah.11615
40. Zavinon F, Adoukonou-Sagbadja H, Bossikponnon A, Dossa H, Ahanhanzo C. Phenotypic diversity for agro-morphological traits in pigeon pea landraces (*Cajanus cajan* L.) Millsp.] cultivated in southern Benin. Open Agric. 2019;4(1):487-99. https://doi.org/10.1515/opag-2019-0046
41. Kumara BN, Santoshagowda GB, Nishanth GK, Dharma Raj PS. Genetic diversity, variability and correlation studies in advanced genotypes of pigeon pea (*Cajanus cajan* L. Millsp.) Genotypes. Acta Biol Indica. 2013;2(2):406-11.
42. Alam MS, Hossain S, Ali MA, Hossain MG, Islam MF. Assessment of genetic divergence in tomato (*Solanum lycopersicum* L.) through clustering and principal component analysis. J Agric Sci Eng Innov. 2020;1(1):10-4. https://doi.org/10.5281/zenodo.3965945

Additional information

Peer review: Publisher thanks Sectional Editor and the other anonymous reviewers for their contribution to the peer review of this work.

Reprints & permissions information is available at https://horizonepublishing.com/journals/index.php/PST/open_access_policy

Publisher's Note: Horizon e-Publishing Group remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Indexing: Plant Science Today, published by Horizon e-Publishing Group, is covered by Scopus, Web of Science, BIOSIS Previews, Clarivate Analytics, NAAS, UGC Care, etc

See https://horizonepublishing.com/journals/index.php/PST/indexing_abstracting

Publisher information: Plant Science Today is published by HORIZON e-Publishing Group with support from Empirion Publishers Private Limited, Thiruvananthapuram, India.