



Phenotypic divergence for agro-morphological traits among extant rice (*Oryza sativa*) varieties

ARCHANA SANYAL¹, MONIKA A JOSHI², B S TOMAR³ and DIVYA AGGARWAL⁴

ICAR- Indian Agricultural Research Institute, New Delhi 110 012

Received: 20 August 2015; Accepted: 2 March 2016

ABSTRACT

The genetic divergence of 61 extant varieties of rice (*Oryza sativa* L.) was assessed by using principal component analysis (PCA). The variables considered in the multivariate analyses were morphological, phenological and grain traits. The correlation analysis revealed that a number of traits were directly associated with other traits regardless of plant type or architectural configuration of the whole plant. The first four components in the PC analysis with Eigen values >1 contributed 74.25% of the variability among genotypes evaluated. Across various PCs, each of the accessions contributed both vegetatively and reproductively and in some cases only one-sidedly, either reproductively or vegetatively. The study grouped the genotypes into seven clusters. Cluster I had maximum number of genotypes (17). The cluster VII had second highest number of genotypes (12), while cluster II, III, IV, V, VI contained nine, four, six, five and eight genotypes, respectively. The overall composition of the clustering pattern showed that genotypes collected from the same geographic origin were distributed in different clusters. A critical appraisal of the observations suggested that none of the clusters contained genotypes with all the desirable traits, which could be directly selected and utilized. Hence, depending on the *per se* performance of the best genotypes within the clusters, they may be directly used for adaptation or maybe used as parents in future breeding programs.

Key words: Cluster analysis, Correlation, Multivariate analysis, Principal component analysis, Rice

India is endowed with a great diversity of rice (*Oryza sativa* L.) germplasm in its vast territorial land area with a variety of short, slender, aromatic rice varieties which are popular in different traditional rice growing pockets. Morphological characterization is an important step in classification of crop germplasm because a breeding program mainly depends upon the magnitude of genetic variability in that species (Tarika *et al.* 2009). Genetic variability studies are hence, important in selection of parents for hybridization as sound crop improvement depends upon the magnitude of variability in the base population. Multivariate statistical methods have found extensive use in summarizing and describing the inherent variation in a population of crop genotypes. Some of the methods include principal component analysis (PCA), discriminant canonical analysis (DCA) and cluster analysis (CA). These statistical techniques identify plant characters that contribute most to the variation within a group of entries. The methods are often extended to genotype grouping in order to cluster entries that show

similarity in one or more characters and thus guide in the choice of parents for hybridization. Among them Principal Component Analysis (PCA) is used to reveal patterns and eliminate redundancy in data sets as morphological and physiological variations routinely occur in crop species. This technique identifies plant traits that characterize the distinctness among selected genotypes. These are often extended to the classification of a population into groups of distinct orders based on similarities in one or more characters, and thus guide in the choice of parents for hybridization. The PCA has been used to partition observed agronomic variations into genotypes of many crops, such as rubber (Omokhae and Alike 2000), rice (Nassir 2002) and sesame (Mponda *et al.* 1997). Various workers studied the relationships between different traits in different crops, such as soybean (Adebisi *et al.* 2001), cassava (Verma and Rai 1993), rice (Chakravorty *et al.* 2013). While PCA helps to identify the traits with the highest variability, correlations reveal the strength of relationships of the identified traits with yield and other traits, which makes the development of new varieties efficient and effective. Correlation studies thus enable the breeder to understand the mutual component characters on which selection can be based for genetic improvement. Keeping in view these facts, the present investigation aimed to characterize morphologically similar looking rice varieties so as to ascertain their variability and

¹Ph D scholar (e mail: archana0811bhu@gmail.com), ² Senior Scientist (e mail: monikakshat622@yahoo.com), Division of Seed Science and Technology ³ Principal Scientist (e mail: bst_spu_iari@rediffmail.com), Seed Production Unit ⁴ Senior Research Fellow (e mail: divyaaggrwl1989@gmail.com). Division of Seed Science and Technology.

diversity.

MATERIALS AND METHODS

The experimental material consisted of 61 extant varieties of rice belonging to both non-basmati and basmati group. The genotypes were raised during *kharif* season 2013 and 2014 at the Seed Technology farm, IARI, New Delhi in randomized complete block design with plot size of 4 rows with each row of 4 meter length. Row to row and plant to plant spacing was maintained at 20 × 15 cm. The material was replicated thrice and all the agronomic practices were followed to raise a good crop. Ten competitive plants were randomly selected from each genotype in each replication to record the data. National DUS Test Guidelines (2003) were followed beginning from the trial layout to recording of the last field-related observation.

As per the guidelines, 46 morphological characters were recorded in rice material at different stages of plant growth. Out of these 46 traits, 12 agro-morphological traits, which were quantitative in nature, were used for diversity analysis in the present study. The variables thus considered in the descriptive and multivariate analyses were morphological (leaf length, leaf width, stem length, panicle length and panicle number), phenological (time of heading) and grain traits (grain weight, length and width of both grain and decorticated grain and L/B ratio).

The correlation between different traits was worked out using Pearson's correlation coefficient (r) which is a measure of strength of linear association between two variables. Multivariate tool used for analyzing data was principal component analysis (PCA). The PCA was calculated by SPAR 1 (Statistical Package for Agricultural Research, developed by IASRI, 1991). Those principal components with Eigen value ≥ 1 were selected for further analysis. An Eigen value represents the amount of variance that is accounted for by a given component. The clustering pattern was assessed using the first two principal components of PCA.

RESULTS AND DISCUSSION

Correlation studies

Correlation pattern among agro-morphological traits is fundamental to provide information for plant breeding program since it helps the breeder to select important characters. The correlation results are presented in Table 1. The positive correlation of stem length with leaf length ($r = 0.438$) indicates the principal of morphogenic compatibility in the architectural plant system of rice. Time of heading was positively correlated with panicle number ($r = 0.432$) which means late flowering provides more number of panicle, ultimately increasing yield. While other traits under study were found negatively correlated with panicle number, viz. leaf length ($r = -0.311$), leaf width, panicle length and stem length. Hence, time of heading could be used in breeding program as yield trait circuitously. It was found that grain width was positively correlated with 1 000 grain weight, indicative of the fact that higher the grain weight, broader be will grain. Along with that, grain length was highly correlated with decorticated grain length and L/B ratio with maximum positive value $r = 0.793$ and $r = 0.875$, respectively. These grain traits affect the grain vigour, because higher the grain size (length, width and L/B ratio), higher will be the vigour (Ries and Everson 1973). The study revealed positive correlation between panicle length, a yield trait with grain length, which also reported by Zafar *et al.* 2006. Traits such as panicle length, grain length and L/B ratio were positively correlated with stem length; suggesting that plant height contributes to yield traits. Leaf width was found to be positively correlated with stem length (0.310), panicle length (0.257) and grain length (0.259) while it was negatively correlated with panicle number. Grain length and decorticated grain length were negatively correlated to grain width and decorticated grain width, thereby showing inverse relationship between length and width, which ultimately affects the L/B ratio. Thus based on the correlation analysis, a number of traits were directly associated with other traits

Table 1 Pearson's correlation coefficient among 12 agro-morphological traits

	LL	LW	TOH	STL	PL	PN	GWT	GL	GW	DGL	DGW	LBR
LL	1											
LW	.364**	1										
TOH	-.077	-.083	1									
STL	.438**	.310*	.074	1								
PL	.340**	.257*	-.238	.522**	1							
PN	-.311*	-.317*	.432**	-.264*	-.522**	1						
GWT	.012	.172	-.190	-.024	.097	-.065	1					
GL	.119	.259*	-.232	.252*	.353**	-.171	.202	1				
GW	-.195	.017	.120	-.199	-.161	.075	.300*	-.473**	1			
DGL	.072	.214	-.114	.128	.233	-.040	.199	.793**	-.349**	1		
DGW	-.126	-.127	.080	-.132	-.260*	.095	.192	-.489**	.715**	-.480**	1	
LBR	.186	.171	-.205	.257*	.330**	-.155	-.044	.875**	-.812**	.739**	-.725**	1

** , Correlation is significant at the 0.01 level ($P < 0.01$), * , Correlation is significant at the 0.05 level ($P < 0.05$). LL; leaf length, LW; leaf width, TOH; time of heading, STL; stem length, PL; panicle length, PN; panicle number, GWT; 1 000 grain weight, GL; grain length, GW; grain width, DGL; decorticated grain length, DGW; decorticated grain width.

regardless of plant type or architectural configuration of the whole plant.

Principal Component Analysis

As it was observed that the characters are interrelated, so to have an idea about their independent impact, principal component analysis was undertaken. The result of the PCA explained the genetic diversity of 61 rice germplasm collection. The per cent variation explained by the four (1-4) most informative principal components and the vector loading for each agro-morphological trait was significant. Those PCs with Eigen values greater than one were selected as proposed by Jeffers (1967). The first four components in the PC analysis with Eigen values >1 contributed 74.25% of the variability among genotypes evaluated for different agro-morphological traits (Table 2). Other PCs (5-12) had Eigen values less than 1, and hence considered as non-significant. Principal component 1 (PC₁), with Eigen value of 4.1708, contributed 34.76% of the total variability, whereas, PC₂ with Eigen value of 2.0785 accounted for 17.32% of the total variability observed among 61 extant rice varieties. While PC₃ and PC₄ had Eigen value of 1.48105 and 1.1799, respectively, and contributed 12.34% and 9.83% of total variability. For each principal axis, there are a number of characters contributing to the total variation (Table 2). In PC₁, the traits that accounted for most of the observed variability among 61 genotypes included grain length with positive loading of 0.4126; decorticated grain length 0.3550 and length/width ratio having value of 0.4523. Decorticated grain width and grain width had highest negative loading with value of -0.3536 and -0.3435, respectively.

The PC₂ explained an additional 17.32% of the total variance. Likewise, in PC₂, the important agro-morphological traits with greater allowance were: length of leaf blade

(0.3460), leaf blade width (0.3445) and panicle length (0.3576) whereas panicle number/plant had high negative loading (-0.4346). The third PC (12.34% of total variance) was associated with high positive loadings of 1 000 grain weight (0.6194); while high negative loading associated with leaf blade length (-0.3065), stem length (-0.3441) and time of heading (-0.3305). The fourth PC explained an additional 9.83% of the total variance with highly positive loadings of time of heading (0.6458), panicle number/plant (0.4241) along with the negative loadings of panicle length of main axis (-0.0842).

Thus, the prominent characters coming together in different principal components and contributing towards explaining the variability have the tendency to remain together which may be kept into consideration during exploitation of these characters in breeding program. Three characters viz. L/B ratio grain, grain length and decorticated grain length contributed highly to PC₁; thus this component is the weighted average of the characters which are both vegetative and reproductive in nature. The characters with higher positive weight in PC₂ were length of leaf blade, width of leaf blade and panicle length; which were vegetative in nature. For PC₃, the character with higher positive loading was 1 000 grain weight which is a reproductive character. For PC₄, the characters with high positive loading were time of heading and panicle number; this component is the weighted average of characters which are both vegetative and reproductive in nature. The findings thus suggest that across various PCs, each of the accessions contribute both vegetatively and reproductively and in some cases, each accession contributes only one-sidedly, either reproductively or vegetatively. PCA is thus, a technique which identifies plant traits that contribute most of the observed variation within a group of genotypes. The tool has a practical application in the selection of parent lines for breeding purposes. The cumulative variance of 74.25% by the first four axes with Eigen value of >1.0 indicates that the identified traits within the axes exhibited great influence on the phenotype of the landraces, and could be effectively used for selection among them.

Cluster analysis

Sixty one extant varieties of rice were grouped into seven clusters Fig 1. The landraces were not evenly distributed among the clusters, and their places of origin were distributed across the clusters. For example, there were 17 landraces in Cluster I, 12 in Cluster VII. These two clusters contained 62 % of the landraces in the study. Cluster I had maximum number of genotypes (17). The cluster VII had second highest number of genotypes (12), while cluster III, IV, V VI contained nine, four, six, five and eight number of genotypes, respectively (Table 3). The overall composition of the clustering pattern showed that genotypes collected from the same geographic origin were distributed in different clusters. Similar findings were also reported by Joshi *et al* (2008). Hence, the kind of genetic diversity found among the genotypes belonging to the same geographic origin

Table 2 Principal component analysis of 12 agro-morphological traits of 61 rice varieties

Variables	PC ₁	PC ₂	PC ₃	PC ₄
DGL	0.355090	-0.174781	0.324033	0.248496
DGW	-0.353678	0.269048	0.159186	0.097713
LL	0.189514	0.346011	-0.306553	0.135654
LW	0.184439	0.344531	0.042432	0.310361
PL	0.280354	0.357642	-0.088142	-0.084242
PN	-0.189639	-0.434674	0.012625	0.424153
SL	0.412655	-0.110169	0.278861	0.137648
STL	0.222733	0.310372	-0.344182	0.310328
SW	0.018565	0.231237	0.619410	0.240114
SWT	-0.343537	0.303678	0.275210	0.192192
TOH	-0.146368	-0.183484	-0.330522	0.645835
L/B ratio	0.452380	-0.232305	0.032870	-0.005309
Eigen Value	4.17081879	2.07859471	1.4810528	1.179912
% variance contribution	34.76	17.32	12.34	9.83
Cumulative variance contribution	34.76	52.08	64.42	74.25

Table 3 *Per se* performance of agro-morphological characters of 61 rice varieties under different clusters

Geno type no.	Genotype	Ch1	Ch2	Ch3	Ch4	Ch5	Ch6	Ch7	Ch8	Ch9	Ch10	Ch11	Ch12
17	VIKRAMARYA, JAYA, PR 113, PANT DHAN 11, KRANTHI, JYOTHI, PANT DHAN 4, RAVI, MANDYA VIJYA, VL DHAN 206, SURAKSHA, VIVEK DHAN 62, HEERA, MAHAMAY, PNR 381, VL DHAN 221, PNR 519,	41.16	1.19	94.05	67.71	24.92	19.76	26.31	8.31	2.58	6.32	2	3.25
9	MAKOM, IR8, LOCHIT, ARUNA, RASI, JD 13, JD 6, ANNADA, CSR 10	32.68	1.20	95.11	48.55	21.21	25.55	24.37	7.91	2.60	5.62	2.06	2.96
4	SALIVAHANA, HMT, PKV SARASWATI, SAMHA-MAHSURI	35.85	1.12	113.5	56	22.12	20.50	13.54	7.42	2.23	5.07	1.85	3.38
6	CSR27, KASTURI, NDR 359 VASUMATI, BASMATI 370, PANT DHAN 12	48.79	1.52	94.33	90.42	27.83	15.33	23.42	9.3	2.01	6.88	1.65	4.63
5	PB 1, IMP.PB1, PUSA-SUGANDHA 2, PUSA-SUGANDHA 3, PUSA- SGANDHA 5	39.32	1.46	70.25	72.43	28.45	14	27.29	11.4	2.00	8.07	1.51	5.70
8	TARAORI BASMATI, VIKASH, GOVIND, SUGANDHAMATI, PNR546, PNR162, PUSA-BASMATI 1121, PUSA BASMATI 1401	38.84	1.25	97.66	66.36	23.89	24.44	21.67	9.95	2.01	7.81	1.50	4.95
12	NIDHI, ASD20, POORNIMA, SHAYMLA, PR106, PUSA 221, KRISHN AVENI, CSR13, PUSA 44, PUSA 33, SABARMATI, PUSA 834	35.84	1.01	110.75	60.66	22.20	25.50	23.11	8.69	2.09	6.66	1.72	4.17

Ch1: Leaf length of blade; ch2: leaf width of blade; ch3: time of heading; ch4: stem length (excluding panicle); ch5: panicle length of main axis; ch6: panicle no. per plant; ch7: grain weight of 1 000 seed; ch8: grain length; ch9: grain width; ch10: decorticated grain length; ch11: decorticated grain width; ch12: L/B ratio.

might be due to change in selection for different objectives, utilizing new germplasm and adopting different methodology for evolving rice varieties. Chakravorty *et al.* (2013) analyzed the diversity of phenotypic traits of landraces of rice and reported that cluster analysis permitted the separation of landraces into 10 major clusters from diverse geographical locations, suggesting the environmental adaptation of landraces.

Analysis of variance indicated that the difference among landraces for all the characters under study in various clusters was significant. The mean, standard error and coefficient of variation for various traits is presented in Table 4.

Clustering based on phenology and morphological characteristics

Cluster II could be characterized by genotypes with maximum panicle number (25.5) and minimum in leaf blade length (32.68 cm), stem length (48.55 cm), panicle length (21.2 cm). Whereas, cluster III represented maximum value

for days to heading (113.5 days; very late flowering). Maximum value of length of leaf blade (48.79 cm), width of leaf blade (1.52 cm) and stem length (90.42 cm) were the characteristics of cluster IV. Cluster VII included genotypes with shortest leaf blade width (1.01 cm). Cluster V could be the descriptive of highest value for panicle length (28.45 cm) along with minimum value of days to heading (70.25 days; early flowering) and panicle numbers/plant (14).

Clustering based on grain characteristics

Clustering of genotypes on the basis of grain characters could be elaborated as following: cluster II represented genotypes with broadest grain width (2.60 cm) as well as decorticated grain width (2.06 cm), along with that it was also a group of genotypes containing lowest L/B ratio of grain (2.96). Cluster III could be identified with lightest 1 000 grain weight (13.54 g), shortest grain length (7.42 cm), and minimum length of decorticated grain length (5.07 cm). Cluster V was representative of highest value of L/B

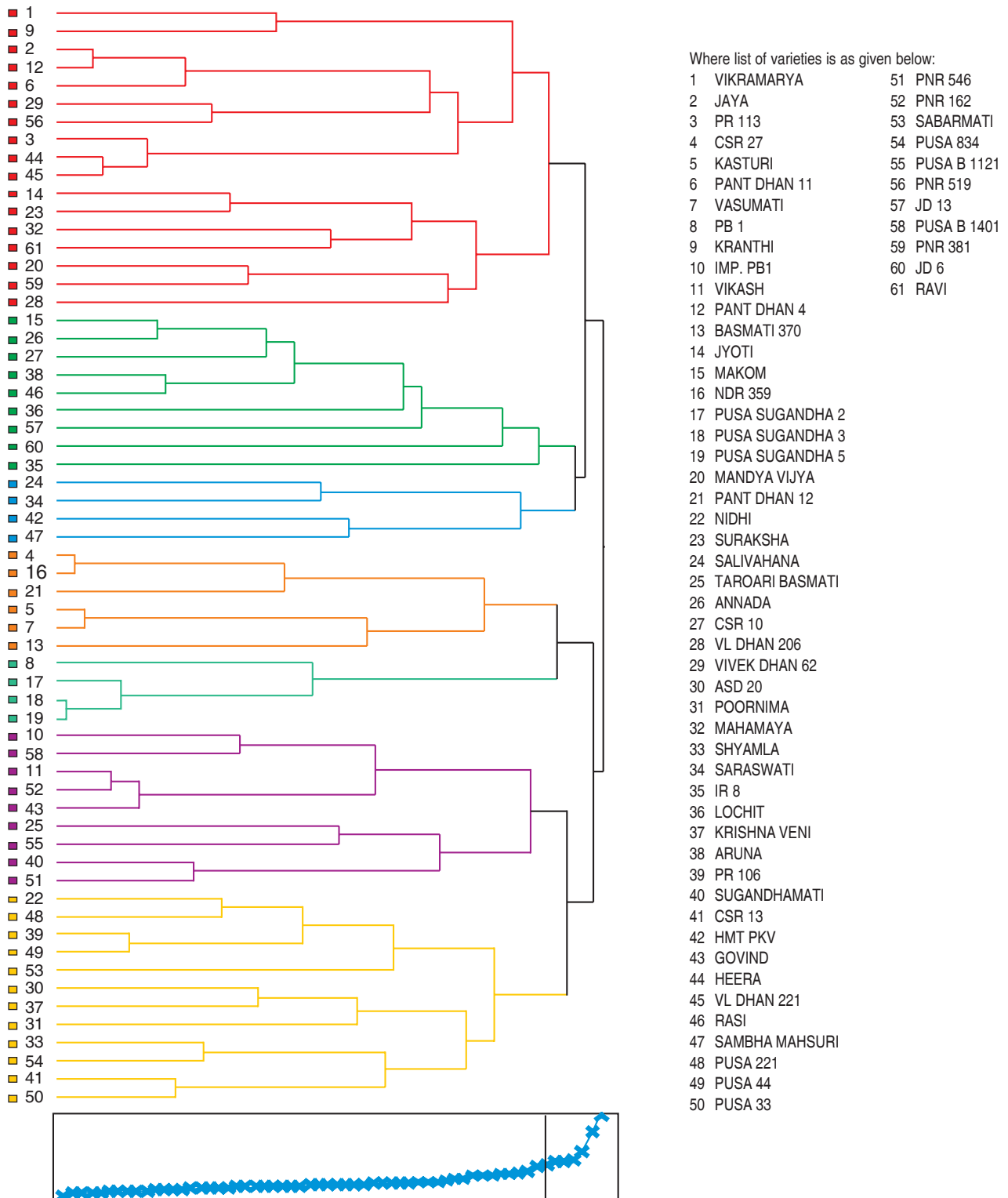


Fig 1 Dendrogram

ratio of grain (5.70), longest grain length (11.4 cm) and longest decorticated grain length (8.07 cm), along with shortest grain width (2.0 cm). Cluster VI genotypes had shortest decorticated grain width (1.50 cm). Further, genotype Ravi had longest leaf blade (55.95 cm) in cluster I while genotype Jaya in the same cluster had smallest number of panicle (12). Likewise, VL Dhan 20 had highest value for

days to heading. Lochit in cluster II had shortest stem length (35 cm) with highest number of panicle in IR 8. In addition to this, four more genotypes of the same cluster recorded maximum and minimum values for various traits, viz. Rasi had longest leaf blade (40 cm); JD 13 had shortest leaf width (0.7 cm); IR8 with longest day to heading (120 days); Makom had lowest L/B ratio (2.5). In cluster III, HMT PKV

Table 4 Means and variation for various traits in respective clusters

Cluster	Character 1 Mean	Character 2 Mean	Character 3 Mean	Character 4 Mean	Character 5 Mean	Character 6 Mean	Character 7 Mean	Character 8 Mean	Character 9 Mean	Character 10 Mean	Character 11 Mean	Character 12 Mean
1	41.067	1.167	94.207	67.307	24.473	19.387	26.09	8.207	2.577	6.343	2.05	3.24
2	32.877	1.2	95.037	48.637	22.063	25.367	24.537	7.333	2.32	5.497	2.053	2.647
3	36.13	1.15	113.157	57.07	22.673	20.883	13.05	7.257	2.31	5.083	1.893	3.28
4	48.247	1.49	94.277	90.263	27.38	15.377	23.16	9.083	2.063	6.84	1.65	4.693
5	39.093	1.47	70.383	72.52	28.337	13.983	27.407	11.163	2.067	8.227	1.503	5.7
6	38.38	1.25	98.06	66.717	23.883	23.957	21.647	9.62	2.013	7.587	1.503	4.54
7	35.147	1.08	110.55	60.59	21.85	25.433	23.043	8.607	2.1	6.67	1.717	4.087
CD	0.821	0.092	0.476	1.592	1.825	0.856	0.669	0.331	0.259	0.418	0.206	0.458
SE(m)	0.264	0.03	0.153	0.511	0.586	0.275	0.215	0.106	0.083	0.134	0.066	0.147
SE(d)	0.373	0.042	0.216	0.723	0.829	0.389	0.304	0.15	0.118	0.19	0.093	0.208
CV	1.180	4.083	0.274	1.338	4.163	2.307	1.638	2.102	6.531	3.519	6.472	6.328

Ch1: Leaf length of blade; ch2: leaf width of blade; ch3: time of heading; ch4: stem length (excluding panicle); ch5: panicle length of main axis; ch6: panicle no. per plant; ch7: grain weight of 1 000 seed; ch8: grain length; ch9: grain width; ch10: decorticated grain length; ch11: decorticated grain width; ch12: L/B ratio.

had longest grain (8.2 cm), heaviest grain weight (16.71 g) in same cluster. In cluster V, Sugandha 3 had highest panicle length (27.67 cm), in contrast with PB 1 with shortest grain (10.4 cm). Similarly, genotype Pusa Basmati 1 121 had longest grain in contrast with broadest grain in Govind (2.1 cm). Along with that, in cluster VIII, Pusa 33 was tallest genotype (75 cm) in contrast with highest value for 1 000 grain weight in Sabarmati (27.23 g).

Hence, a critical appraisal of the observations suggested that none of the clusters contained genotypes with all the desirable traits, which could be directly selected and utilized thereby underlining the fact that the hybridization between genotypes of different clusters is necessary for the development of desirable genotypes. Therefore, depending on the *per se* performance of the best genotypes within the clusters, they may be directly used for adaptation or maybe used as parents in future hybridization programs. The study on genetic divergence thus leads to the suggestion that varieties and the process associated with the varietal development are essentially consequences of adaptation of populations to their agro-ecological conditions, aided by utility value (best practice) to the farmer. The agro-morphological characters are thus variable among the landraces, and highlight the difficulty of using them as sole markers for delineating different groups. The clustering pattern can thus help in formulating actual breeding strategies.

REFERENCES

- Adebisi M A, Ariyo O J and Kehlnde O B. 2001. Variation and correlation studies in quantitative characteristics in soybean. *Proceedings of the 35th Annual Conference of the Agricultural Society of Nigeria*, September, pp: 20 held at the University of Agriculture, Abeokuta, 16-121-5.
- Chakravorty A, Ghosh P D and Sahu P K. 2013. Multivariate analysis of phenotypic diversity of landraces of rice of West Bengal. *American Journal of Experimental Agriculture* **3** (1): 11-123.
- Jeffers J N R. 1967. Two case studies in the application of principal component analysis. *Applied Statistics* **16**: 225-36.
- Joshi M A, Singh P, Sarao N K, Sharma R C and Bharaj T S. 2008. Genetic diversity among rice varieties cultivated in Punjab. *Oryza* **45** (4): 277-79.
- Mponda O K K, Gobbon D, Arthur E and Morse S. 1997. Involving farmers in the design of a low input control programme for sesame flea beetles in Southern Tanzania. *Experimental Agriculture* **33** (3): 313-20.
- Nassir A. 2002. Studies on genotype \times environment interactions, variability and plant character correlation in rice. Ph D Thesis, University of Agriculture, Abeokuta, Nigeria, p 123.
- Omokhae K O and Alika J E. 2000. Multivariate of agronomic data on Heveabrsiliensis. (In) *Biometry and quality of life. Nokoe S(Ed)*. Proceedings of SUSAN-IBS Meting, 23-27 August, 1999, Ibadan, Nigeria SUSAN-IBS, Nairobi, Kenya; pp168-173.
- Ries S K and Everson E H. 1973. Protein content and grain size relationships with grainling vigour of wheat cultivars. *Agronomy Journal* **65**: 884-6.
- Tarika Y T, Somta P and Srinives P. 2009. Genetic variation in cultivated mungbean germplasm and its implication in breeding for high yield. *Field Crops Research* **112**: 260-6.
- Verma S P and Rai M. 1993. Genetic variability and inter-relation in cassava (*Manihot esculenta Crantz*) under rainfed conditions of Tripura. *Journal of Root Crops* **19** (2): 77-80.
- Zafar N, Aziz S and Masood S. 2006. Phenotypic divergence for agro-morphological traits among landrace genotypes of rice (*Oryza sativa* L.) from Pakistan. *International Journal of Biological Science* **6** (2): 335.