The Andhra Agric. J 55(4):442-450, 2008

Principal Component and Cluster analysis in Chickpea (*Cicer arietinum* L.)

U Lakshmi Annapurna, S Srimannarayana Murthy, C Panduranga Rao and R Srinivasulu
Department of Genetics & Plant Breeding, Agricultural College, Bapatla 522 101, Andhra Pradesh.

ABSTRACT

Seventy genotypes of chickpea were evaluated to study genetic divergence by using principal component and cluster analysis. These genotypes were grouped into 9 clusters. Principal components with eigen values more than the one contributed 89.34 per cent of the cumulative variance. Higher inter cluster distance was observed between cluster IV and IX followed by cluster IV and VI. In hierarchical cluster analysis the clustering pattern of genotypes was to be independent of their eco-geographical origin.

Key words : Chickpea, Cluster Analysis,Genetic Divergence, Principal Component Analysis.

The success of any breeding programme depends upon the availability of adequate genetic diversity. The major factor responsible for limited success in increasing the chickpea yield has been the narrow genetic base of the material available. Genetic diversity plays an important role, because hybrids between lines of diverse origin generally, display a greater heterosis than those between closely related parents. Genetic diversity is the basic criterion for the continuous improvement of the crop whether through natural section (or) direct plant breeding. The present study aims at assessment of genetic divergence in 70 genotypes of Bengalgram through Hierarchical cluster analysis and Principal Components Analysis (PCA). Crosses can be made between these genotypes for character improvement and seed yield.

MATERIAL AND METHODS

Seventy genotypes of chickpea (Table 1) were sown in the *rabi* season of 2006-07 in a randomized block design with three replications at Regional Agricultural Research Station, Lam, Guntur. Each entry was planted in a single row of 4 m length with a spacing of 30 x 10 cm. The observations were recorded on ten randomly selected plants in each entry and in each replication on 11 component characters i.e. days to 50% flowering, days to maturity, plant height (cm), number of primary branches per plant, number of secondary branches per plant, number of pods per plant, 100-seed weight (g), harvest index (%), biological yield per plant (g), protein content (%) and seed yield per plant (g) and mean values were used for statistical analysis. The data were analyzed using hierarchical cluster analysis (Anderberg, 1993) and principal component analysis (Jackson, 1991).

RESULTS AND DISCUSSION

The analysis of variance revealed highly significant differences among the 70 genotypes of Bengalgram indicating that the existence of substantial genetic variability for all the characters under study.

Principal component analysis (PCA) identified four principal components with eigen values more than one which contributed 89.3 per cent of cumulative variance (Table 2). The first principal component (PC_1) contributed maximum towards variability (63.12%) with significant loading of 100-seed weight (-0.8) and biological yield per plant, (-0.385) which were negatively correlated and number of primary branches per plant (0.056) which was positively correlated. The second principal component (PC_2) accounted for 10.24 per cent of total variance and it reflected significant loading of days to 50% flowering (0.772) and plant height (0.289) which were positively correlated. The third principal component (PC_3) accounted for 9.05 per cent of cumulative variance and it was characterized by conspicuously high loading for days to maturity (-0.610), and number of secondary branches per plant (-0.375), which were negatively correlated and number of pods per plant (0.36) which was positively correlated (Table 3).Based on these first three principal components mean genotype scores were computed (Table 4). Principal factors scores for all the 70 genotypes were estimated in all 3 PC's and utilized to construct precise 2D and 3D plot (Fig 1 and 2). All the genotypes were plotted for PC_1, PC_2 and PC_3 which cumulatively explained 82.43 per cent of variability accounted for all the characters.

The plot of PC_1, PC_2 and PC_3 showed characters differentiation of genotypes according to their cluster membership for each cluster. The mean

Table 1. List of 70 chickpea (*Cicer arietinum* L.) genotypes and their source of origin

Sl.No.	Genotype / germplasm accession No.	Source
1	GNG 469	ARS, Sriganganagar
2	BG 256	IARI, New Delhi
3	Pusa 372	IARI, New Delhi
4	JG 315	JNKVV, Jabalpur
5	L 550	PAU, Ludhiana
6	KWR 108	CSAUAT, Kanpur
7	RSG 888	ARS, Durgapura
8	SAKI 9516	JNKVV, Jabalpur
9	HC -3	HAU, Hisar
10	Avrodhi	GBPUAT, Pantnagar
11	Pant G 186	GBPUAT, Pantnagar
12	C 235	PAU, Ludhiana
13	BGM 408	IARI, New Delhi
14	Radhey	GBPUAT, Pantnagar
15	DCP 92-3	GBPUAT, Pantnagar
16	K-850	GBPUAT, Pantnagar
17	Dahood yellow	ARS, Junagadh
18	ICC 3219	ICRISAT, Hyderabad
19	ICC 92338	ICRISAT, Hyderabad
20	ICC 7425	IIPR, Kanpur
21	ICC 5168	PAU, Ludhiana
22	ICC 4948	PAU, Ludhiana
23	ICC 3500	ICRISAT, Hyderabad, (collection by RPIP, Turke from Ankara)
24	ICC 432	ICRISAT, Hyderabad, (collection by RPIP, Pusa, Bihar)
25	ICC 12237	ICRISAT, Hyderabad
26	ICC 453	ICRISAT, Hyderabad, (collection by RPIP, Pusa, Bihar)
27	ICC 506	ICRISAT, Hyderabad, (collection by ANGRAU, Rajendranagar)
28	ICC 3296	ICRISAT, Hyderabad, (collection from Iran)
29	ICC 1564	ICRISAT, Hyderabad, (collection from IIPR, Kanpur)
30	ICC 16644	ICRISAT, Hyderabad, (collected from Pakistan)
31	ICC 12373	ICRISAT, Hyderabad, (collected from Maharashtra)
32	JG 11I	CRISAT, Hyderabad,
33	ICCV-10	ICRISAT, Hyderabad
34	ICCC 37	ICRISAT, Hyderabad
35	ICC 927	ICRISAT, Hyderabad, (collection from IARI, New Delhi)

Sl.No.	Genotype / germplasm accession No.	Source
36	ICC 706	ICRISAT, Hyderabad (collected by IARI from Punjab)
37	Jyothi	RARS, Lam
38	Vijay	MPKV, Rahuri
39	Annegiri	UAS, Dharwad
40	ICC 14694	ICRISAT, Hyderabad (collected by JNKVV, Jabalpur)
41	Pusa 1053	IARI, New Delhi
42	HK 3	HAU, Hisar
43	ICC 12331	ICRISAT, Hyderabad (collected by IARI from Syria)
44	ICC 4908	ICRISAT, Hyderabad (collected by RPIP from Turkey)
45	ICCV 89224	ICRISAT, Hyderabad
46	ICC 5320	ICRISAT, Hyderabad developed by PAU, Ludhiana
47	IC 12495	ICRISAT, Hyderabad (collected by ICRI, India)
48	JGK -1	JNKVV, Jabalpur
49	ICC 4929	ICRISAT, Hyderabad (collected by PAU, Ludhiana)
50	ICCV 04312	ICRISAT, Hyderabad
51	ICCV 03407	ICRISAT, Hyderabad
52	ICCV 04301	ICRISAT, Hyderabad
53	ICCV 04309	ICRISAT, Hyderabad
54	ICCV 04305	ICRISAT, Hyderabad
55	ICCV 04308	ICRISAT, Hyderabad
56	ICCV 04311	ICRISAT, Hyderabad
57	ICCV 05306	ICRISAT, Hyderabad
58	ICCV 05310	ICRISAT, Hyderabad
59	ICCV 05311	ICRISAT, Hyderabad
60	ICCV 05313	ICRISAT, Hyderabad
61	ICCV 05314	ICRISAT, Hyderabad
62	ICCV 05315	ICRISAT, Hyderabad
63	Virat	ICRISAT, Hyderabad
64	Dollar	ICRISAT, Hyderabad
65	ICCV 2	ICRISAT, Hyderabad
66	KAK – 2	MPKV, Rahuri
67	Vihar	MPKV, Rahuri
68	ICCV 95334	ICRISAT, Hyderabad
69	LBeG 7	RARS, Lam
70	JGK-2	JNKVV, Jabalpur

Table 2. The eigen values, per cent variance, cumulative percent variance for four principal components in chickpea (*Cicer arietinum* L.)

Character	PC_1	PC_2	PC_3	PC_4
Eigen value	59.50	23.96	22.50	19.69
% Of variance	63.12	10.24	9.05	6.91
Cumulative variance	63.12	73.37	82.43	89.34

Table 3. Character loading of four principal components for 70 different genotypes of chickpea (*Cicer arietinum* L.)

Character	PC_1	PC_2	PC_3	PC_4
Days to 50% flowering	0.115	0.722	0.5411	0.135
Days to maturity	-0.068	0.460	-0.610	-0.202
Plant height (cm)	-0.070	0.289	0.090	0.186
No. of Primary branches $plant^{-1}$	0.056	0.009	-0.019	-0.019
No. of Secondary branches $plant^{-1}$	-0.009	0.330	-0.375	0.355
No. of Pods $plant^{-1}$	-0.046	-0.039	0.364	-0.243
100-seed weight (g)	-0.884	0.054	0.138	0.026
Harvest index (%)	0.002	-0.065	0.118	-0.181
Biological yield $plant^{-1}$ (g)	-0.385	0.017	-0.034	-0.096
Protein content (%)	0.091	0.250	0.006	-0.797
Seed yield $plant^{-1}$ (g)	0.178	-0.000	-0.133	-0.202

scores of genotypes were used as input for clustering in order to group the genotypes into various clusters. Hierarchical clustering procedure (Ward's method) was followed to group the genotypes into 9 clusters (Table 5 Fig 3). This reflects that there was no relation between geographical origin and genetic diversity.

The biggest cluster was cluster III consisting of 18 genotypes followed by cluster VIII comprising of 14 genotypes. Based on cluster analysis the intra cluster values ranged from 0.00 (Cluster IV) to 195.41 (Cluster VI), the maximum inter cluster distance was observed between cluster IV and cluster IX (2084.07) followed by cluster IV and VI (1951.49) and cluster IV and VIII (1577.04) as shown in (Table 6). Cluster IV is characterized by high mean value for days to 50% flowering, days to maturity, number of primary branches per plant, protein content (%), cluster VII recorded high mean values for plant height (cm) and number of secondary branches per plant. Cluster IX is characterized by high mean values for no. of pods per plant, 100-seed weight, harvest index, biological yield per plant (Table 7). Based on these studies crosses may be effective between the genotypes of these clusters to obtain better and desirable segregants. Utilization of principal component analysis combined with hierarchical cluster analysis in genetic diversity studies was reported by earlier workers Ghafoor *et al.* (2001) in black gram and Narendra Singh in bengalgram (2002).

The present study depicted the relative divergence in morphological and yield traits. The clustering pattern could be utilized in identifying the best cross combinations for generating variability with respect to various traits under study. The genotypes clubbed in the different clusters if inter crossed may generate wide variability. Some combinations may also exhibit high heterosis for seed yield and transgressive segregants for yield and yield components may also be expected.

Table 4. The PCA scores or genotypic mean scores for 70 genotypes of chickpea

Sl. No.	Geno type	X- vector	Y- vector	Z- vector	Sl. No.	Genotype	X- vector	Y- vector	Z- vector
1	GNG 469	-6.609	35.850	3.674	36	ICC 706	-7.601	32.051	0.497
2	BG 256	-8.411	33.039	4.778	37	Jyothi	-1.408	32.855	-1.130
3	Pusa 372	0.890	33.760	1.347	38	Vijay	0.678	34.698	0.506
4	JG 315	-4.034	35.975	2.214	39	Annegiri	-0.121	32.528	-1.241
5	L 550	-2.122	39.014	7.602	40	ICC 14694	-7.733	33.527	8.585
6	KWR 108	-2.347	37.443	-2.834	41	Pusa 1053	-6.592	37.934	-0.504
7	RSG 888	-0.498	32.800	0.917	42	HK 3	-9.698	39.333	-1.655
8	SAKI 9516	-1.647	35.365	6.482	43	ICC 12331	-13.319	44.030	0.309
9	HC -3	-4.598	35.586	-0.811	44	ICC 4908	-10.714	41.672	-0.484
10	Avrodhi	-0.596	36.214	0.426	45	ICCV 89224	-16.714	30.922	2.410
11	Pant G 186	-1.285	36.909	-0.556	46	ICC 5320	5.885	44.317	0.010
12	C 235	1.842	34.836	2.433	47	IC 12495	-9.245	32.118	-0.592
13	BGM 408	2.419	36.793	-1.986	48	JGK -1	-11.882	33.159	1.633
14	Radhey	-5.160	34.186	0.184	49	ICC 4929	4.150	36.190	0.005
15	DCP 92-3	1.081	35.292	5.346	50	ICCV 04312	-10.761	34.432	1.323
16	K-850	-9.835	34.110	-0.473	51	ICCV 03407	-16.591	35.853	2.031
17	Dahood	-1.275	31.564	3.726	52	ICCV 04301	-15.656	34.532	4.874
18	ICC 3219	-0.758	34.294	4.718	53	ICCV 04309	-18.516	36.082	1.249
19	ICC 92338	-5.271	33.063	-2.036	54	ICCV 04305	-12.251	37.265	0.400
20	ICC 7425	-15.265	29.031	-1.520	55	ICCV 04308	-17.274	34.418	0.299
21	ICC 5168	1.370	32.652	9.381	56	ICCV 04311	-13.627	36.474	0.258
22	ICC 4948	0.890	33.813	-4.036	57	ICCV 05306	-13.271	34.496	0.993
23	ICC 3500	-2.088	32.503	5.711	58	ICCV 05310	-16.975	35.851	3.177
24	ICC 432	-0.229	34.805	1.416	59	ICCV 05311	-15.858	36.863	1.086
25	ICC 12237	1.800	36.177	0.102	60	ICCV 05313	-15.483	34.830	-1.610
26	ICC 453	-0.380	34.188	-2.464	61	ICCV 05314	-13.888	35.534	2.422
27	ICC 506	-1.077	33.359	2.889	62	ICCV 05315	-14.921	36.636	3.313
28	ICC 3296	2.523	33.895	2.940	63	Virat	-12.289	35.568	1.605
29	ICC 1564	1.349	32.761	-2.252	64	Dollar	-23.653	38.501	2.024
30	ICC 16644	-3.064	31.923	1.976	65	ICCV 2	-9.687	27.247	-2.945
31	ICC 12373	-9.905	34.248	2.561	66	KAK – 2	16.373	37.683	1.526
32	JG 11	-4.790	34.010	3.477	67	Vihar	14.379	37.984	1.381
33	ICCV-10	-2.126	34.141	-2.668	68	ICCV 95334	22.442	31.563	0.353
34	ICCC 37	-5.406	34.246	-2.215	69	LBeG 7	12.404	31.102	0.023
35	ICC 927	-12.003	31.093	0.809	70	JGK-2	14.871	33.27	0.841

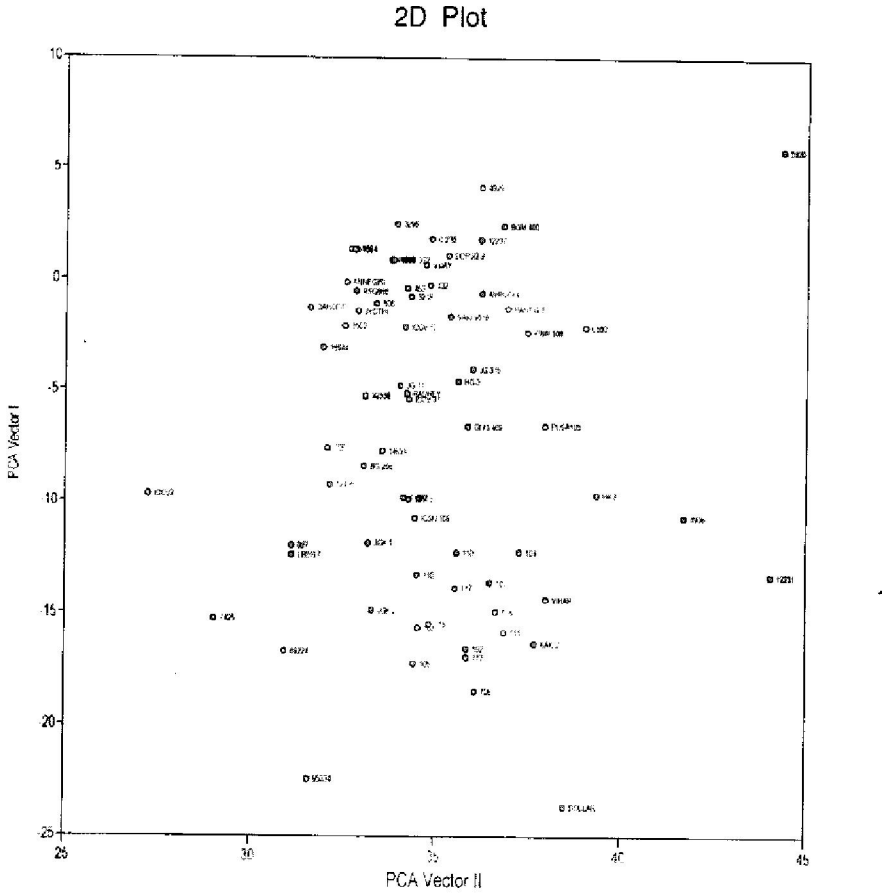

Fig. 1. Two dimensional graph showing relative position of genotypes of chickpea (*Cicer arietinum* L*)* based on PCA scores

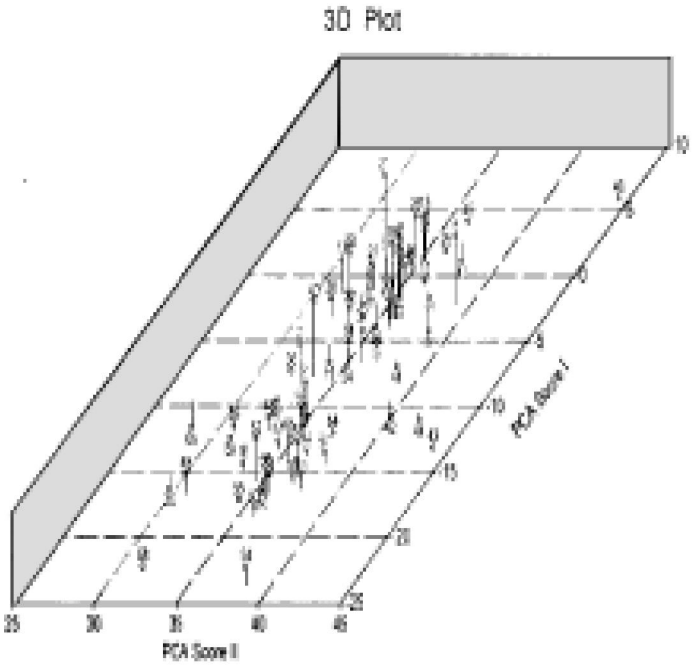

Fig. 2. Three dimensional graph showing relative position of genotypes of chickpea (*Cicer arietinum* L*)* based on PCA scores

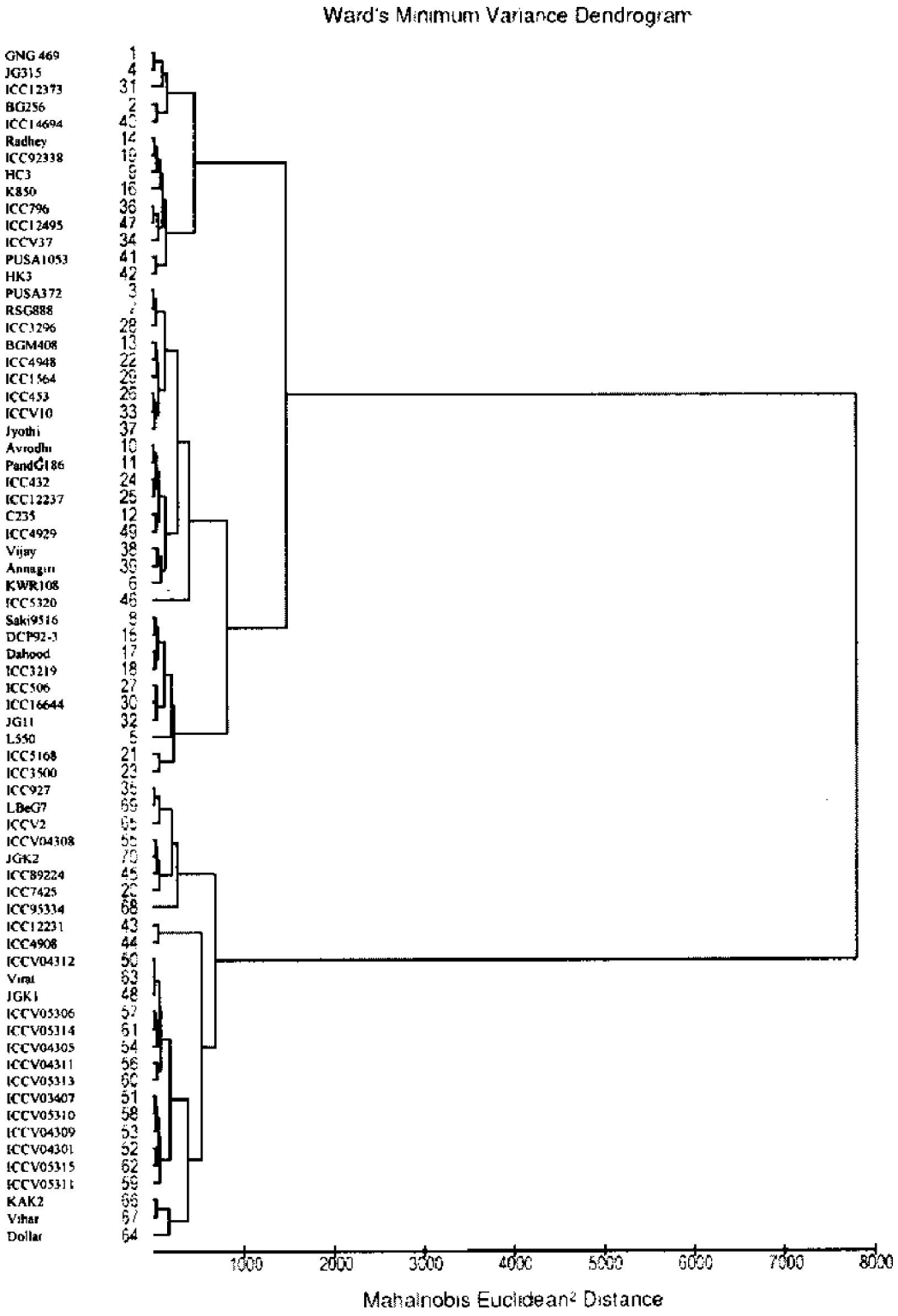

Fig. 3. Hierarchical clustering procedure of the chickpea genotypes using Ward's minimum variance method

Table 5. Cluster position of 70 genotypes of chickpea *(Cicer arietinum* L.) based on Principal Component Analysis

Cluster's	No. of genotypes	Genotypes
I	5	GNG-469, JG-315, ICC-12373, BG-256, ICC-14694
II	9	Radhey, ICC-92338, HC-3, K-850, ICC-706, ICC-12495, ICC-37, Pusa 1053, HK-3
III	18	Pusa 372, RSG 888, ICC-3296, BGH-408, ICC-4948, ICC-1564, ICC-453, ICCV-10, Jyothi, Avrodhi, Pant G 186, ICC-432, ICC-12237, C-235, ICC-4929, Vijay, Annegiri, KWR 108,
IV	1	ICC-5320
V	10	SAKI 9516, DCP 92-3, Dahood yellow, ICC 3219, ICC-506, ICC-16644, JG-11, L-550, ICC-5168, ICC-3500
VI	8	ICC-927, LBeG7, ICCV-2, ICCV-04308, JGK-2, ICC-89224, ICC-7425, ICC-95334
VII	2	ICC-12231, ICC-4908
VIII	14	ICCV-04312, Virat, JGK-1, ICCV-05306, ICCV-05314, ICCv-04305, ICCV-04311, ICCv-05313, ICCV-03407, ICCV-05310, ICCV-04309, ICCV-04301, ICCV-05315, ICCV-05311
IX	3	KAK-2, Vihar, Dollar

Table 6. Mean intra (bold) and inter cluster distance among nine clusters using Ward's minimum variance method in chickpea *(Cicer arietinum* L.)

Cluster	I	II	III	IV	V	VI	VII	VIII	IX
I	**173.59**	273.53	405.04	1033.11	314.08	472.15	610.60	330.02	636.34
II		**135.39**	303.87	877.74	348.03	390.77	416.46	294.86	683.27
III			**128.35**	485.09	260.16	926.18	866.15	783.17	1281.05
IV				**0.00**	666.68	1951.49	1030.34	1577.04	2084.07
V					**171.79**	835.43	813.56	680.10	1171.50
VI						**195.41**	607.52	215.23	477.425
VII							**105.42**	356.93	565.60
VIII								**91.03**	259.90
IX									**190.37**

Table 7. Cluster means for 11 different character in chickpea (*Cicer arietinum* L.)

Characters	Days to 50% flowering	Days of Maturity	Plant height (cm)	No of primary branches plant^{-1}	No. Of. Secondary / branches plant^{-1}	No. Of. Pods plant^{-1}	100 seed weight (g)	Harvest index (%)	Biological yield plant^{-1} (g)	Protein content (%)	Seed yield per plant (g)
Cluster I	55.53	90.73	41.01	2.42	4.46	39.98	26.56	41.72	23.45	21.68	9.79
Cluster II	50.00	92.85	43.89	2.02	11.38	28.94	24.62	38.48	22.38	18.01	8.62
Cluster III	52.24	94.13	39.34	2.12	9.30	30.78	14.15	38.16	15.13	19.21	5.88
Cluster IV	72.33	112.66	43.66	2.96	19.73	46.0	12.83	22.37	24.69	20.71	5.44
Cluster V	57.63	88.96	42.63	1.86	7.58	45.32	18.02	44.35	18.26	18.71	8.17
Cluster VI	42.04	83.62	39.84	1.82	8.58	31.47	35.36	42.66	25.03	18.02	10.37
Cluster VII	63.50	106.50	54.93	2.70	21.91	48.1	37.08	35.03	38.03	19.50	13.50
Cluster VIII	52.59	93.97	45.53	1.66	9.14	38.99	35.06	46.28	34.43	18.77	15.80
Cluster IX	53.667	102.00	48.55	1.56	9.71	76.74	40.87	51.06	53.11	19.90	26.97

LITERATURE CITED

Anderberg M R 1993. Cluster Analysis for Applications. New York: Academic Press.

Ghafoor A,Sharif A, Ahmed Z, Zahid M A and Rabbani M A 2001. Genetic diversity in blackgram (*Vigna mungo* L. Hepper). Field Crops Research, 69: 183-190.

Jackson J E 1991. A User's Guide to Principal Components. John Wiley and Sons, New York.

Narendra Singh 2002. Multivariate analysis in *Kabuli* chickpea accessions. Indian Journal of Pulses Research, 15(1): 70-72.

(Received on 31.10.2007 and revised on 28.01.2008)