Indian J. Plant Genet. Resour. 35(2): 178–184 (2022)
DOI 10.5958/0976-1926.2022.00018.3

RESEARCH ARTICLE

Genetic Divergence Assessment through K-Means Clustering and Principal Component Analysis for Seed Yield, Zinc, Iron and Protein Content in *Vigna unguiculata* L. Walp.

CA Manoj, Marappa*, N Keerthi, A Patil, S Ramesh, DV Naveen, MSP Kanavi, Gangadhar Eshwar Rao, P Venkataravana and DL Savithramma
Department of Genetics and Plant Breeding, UAS, GKVK, Bengaluru-560065, Karnataka, India

(Received: 08 June, 2021; Revised: 22 October, 2021; Accepted: 26 October, 2021)

Cowpea [*Vigna unculata* (L.) Walp.] with chromosome number 2n=24 is an one of the important crops grown in arid and semi-arid regions of the world. Cowpea is a source of high quality protein consumed in many parts of the world. Malnutrition is a global problems needs to be addressed locally. Identification of micronutrient dense genotypes is the major objective of cowpea Biofortification programme. In the present study, 169 cowpea genotypes collected from different sources were analysed for genetic divergence through K-means clustering followed by Principal component analysis (PCA). Considering mean values, genotypes were classified into eight clusters following K-means clustering algorithm. Cluster VII has highest number of genotypes (56 genotypes) followed by cluster III (43 genotypes). Clusters VI and VIII were solitary with single genotype. Results revealed that crosses could be made between genotypes of cluster V & VIII, cluster III & VIII and cluster VII & VIII. According to PCA, Five principal components (PC) with Eigen values greater than one contributed 70.5% of the total variability. Amongst, PC_1 accounted highest proportion of total variance (31.3 %), remaining PCs *viz*., PC_2, PC_3, PC_4 and PC_5 revealed 14.6%, 10.1%, 7.7% and 6.8 % respectively, revealing pod yield^{-1}, seed yield^{-1}, pod length and seeds pod^{-1} were major contributors to the total genetic variability and divergence.

Key Words: Cowpea, Genetic divergence, K-means clustering, Principal Component Analysis (PCA)

Introduction

Cowpea (*Vigna unguiculata* L. Walp) with chromosome number 2n=24 belonging to Leguminaceae family is widely grown in many parts of the world and has multifarious uses as pulse, vegetable, green manure and fodder. It is a one of major sources of calories, protein and minerals in developing countries (Ira *et. al*, 2020). Cowpea dry seeds (per 100g) contains carbohydrates 59.54g, protein 24g, iron 9.95mg, zinc 6.11 mg and low saturated fat 0.5g (https://www.nutritionvalue.org/Cowpeas%2C_raw%2C_mature_seeds%2C_catjang_nutritional_value.html?size=100+g). The focus of green revolution during 1960's was to improve yield levels of major food crops namely, wheat, rice and maize. The dwarf varieties, which occupied the most of cultivation area, helped to fight hunger. Monoculture of these crops resulted in unforeseen rise in malnutrition (Howarth and Ross, 2010). Most essential micronutrients vital to growth and development include iron, zinc whose deficiency can cause serious health problems like stunted growth and anaemia. Though several options are available for development of biofortified cowpea, development of biofortified cowpea via breeding approaches is a economical, viable, environmentally friendly and sustainable option. Identification of micronutrient dense genotypes is an essentially important step for development of biofortified cowpea varieties. This is only possible though a careful selection parents which are genetically divergent. The genetic improvement of cowpea begins with the selection of parents and formation of the base population generating segregating populations in which superior lines are selected, making knowledge on the dissimilarity between the parents particularly important. Several multivariate techniques are helpful to predict genetic divergence such as clustering method and/or principal component analysis. Cluster analysis, a technique of multivariate analysis, groups the plants based on their traits, so that the internal homogeneity and the external heterogeneity of the ensuing groups should be high. Similarly principal component analysis technique

---

*Author for Correspondence: E-mail: nmaars@gmail.com

Table 1. List of cowpea genotypes used in the study and their source

Sl. No.	Genotype	Source	Sl. No.	Genotype	Source	S;. No.	Genotype	Source
1	PV-3-1	NBPGR, Jodhpur	48	EC-472250	NBPGR, Jodhpur	95	ArkaGarima	IIHR, Bengaluru
2	MFC-09-15	VC Farm, Mandya	49	PKB-4-1	GKVK, Bengaluru	96	IIHR-144	IIHR, Bengaluru
3	EC-458480	IIPR,Kanpur	50	MFC-09-10	VC Farm, Mandya	97	NBC-19	NBPGR, Delhi
4	IC-402180	IIPR, Lucknow	51	NBC-8	NBPGR, Delhi	98	EC-458473	IIPR,Kanpur
5	EC-458473	NBPGR, Jodhpur	52	EC-458402	IIPR,Kanpur	99	C-720	IIPR,Kanpur
6	IC-249588	IIPR,Kanpur	53	202804(83)	IIPR,Kanpur	100	NBC-47	NBPGR, Delhi
7	EC-402159	IIPR,Kanpur	54	CB-10	NBPGR, Jodhpur	101	NBC-43	NBPGR, Delhi
8	IC-2591054	IIPR,Kanpur	55	MS-4	IIPR,Kanpur	102	IC-202781	IIPR,Kanpur
9	TOME-774	NBPGR, Jodhpur	56	PGCP-6	Pant Nagar, UP	103	C-720	NBPGR, Jodhpur
10	PV-1-3	NBPGR, Jodhpur	57	MFC-09-17	VC Farm, Mandya	104	CHILORE-11	Shimoga local
11	AV-2-2	GKVK, Bengaluru	58	PGCP-3	Pant Nagar, UP	105	NBC-30	NBPGR, Delhi
12	MFC-09-19	VC Farm, Mandya	59	PGCP-12	Pant Nagar, UP	106	IC-402125	IIPR,Kanpur
13	MFC-09-16	VC Farm, Mandya	60	Pant lob-3	Pant Nagar, UP	107	EC-170584	NBPGR, Jodhpur
14	EC-458505	IIPR,Kanpur	61	Pant lob-2	Pant Nagar, UP	108	NBC-39	NBPGR, Delhi
15	IC-249593	NBPGR, Jodhpur	62	PGCP-5	Pant Nagar, UP	109	EC-402104	IIPR,Kanpur
16	MFC-09-20	VC Farm, Mandya	63	Pantlob-1	Pant Nagar, UP	110	IC-402180	IIPR,Kanpur
17	C-457	IIPR,Kanpur	64	PGCP-27	Pant Nagar, UP	111	EC-472252	IIPR,Kanpur
18	IC-202711(58)	IIPR,Kanpur	65	AV-5	GKVK, Bengaluru	112	IC-402090	NBPGR, Jodhpur
19	IT38956-1	NBPGR, Jodhpur	66	PKB-1	GKVK, Bengaluru	113	C-24-1	IIPR,Kanpur
20	MFC-09-8	VC Farm, Mandya	67	PKB-2	GKVK, Bengaluru	114	MFC-09-05	VC Farm, Mandya
21	IC-21483	IIPR,Kanpur	68	PV-3	NBPGR, Jodhpur	115	IT-97154299	IIPR,Kanpur
22	AV-5-1	GKVK, Bengaluru	69	PKB-4-2	GKVK, Bengaluru	116	IC-1071	IIPR,Kanpur
23	NBC-44	NBPGR, Delhi	70	EC-170584	IIPR,Kanpur	117	V-16	NBPGR, Jodhpur
24	NBC-38	NBPGR, Delhi	71	IIHR-140	IIHR, Bengaluru	118	IC-202867	NBPGR, Jodhpur
25	MFC-09-3	VC Farm, Mandya	72	IC-201	NBPGR, Jodhpur	119	EC-170604	NBPGR, Jodhpur
26	PV-1-4	NBPGR, Jodhpur	73	IIHR-137	IIHR, Bengaluru	120	C-33	Shimoga local
27	IC-402114	IIPR,Kanpur	74	PKB-3	GKVK, Bengaluru	121	MFC-09-02	VC Farm, Mandya
28	V-585	IIPR,Kanpur	75	MBC-25	IIPR,Kanpur	122	IC-202781	NBPGR, Jodhpur
29	IC-402101	IIPR,Kanpur	76	NBC-16	NBPGR, Delhi	123	NBC-24	NBPGR, Delhi
30	MFC-09-7	VC Farm, Mandya	77	PV-1	NBPGR, Jodhpur	124	IC-201	NBPGR, Jodhpur
31	Genotype-36	NBPGR, Jodhpur	78	IC-402106	NBPGR, Jodhpur	125	27749(25)	IIPR,Kanpur
32	PV-3-4	NBPGR, Jodhpur	79	EC-458438	IIPR,Kanpur	126	IC-402101	NBPGR, Jodhpur
33	IC-4506	NBPGR, Jodhpur	80	Kashi Kanchan	NBPGR, Jodhpur	127	GC-3	IIPR,Kanpur
34	IC-402161	IIPR,Kanpur	81	AV-1	GKVK, Bengaluru	128	IC-402180	IIPR,Kanpur
35	IC-4506	IIPR,Kanpur	82	IC-1061	IIPR,Kanpur	129	NC-32	NBPGR, Jodhpur
36	C-24-1	IIPR,Kanpur	83	PKB-5	GKVK, Bengaluru	130	EC-394839	NBPGR, Jodhpur
37	ETC-27	NBPGR, Jodhpur	84	APC-1218	IIPR,Kanpur	131	MBC-25	IIPR,Kanpur
38	EC-458473	IIPR,Kanpur	85	IC-58905	IIPR,Kanpur	132	MFC-09-20	VC Farm, Mandya
39	EC-458489	NBPGR, Jodhpur	86	Bhagyalakshmi	IIPR,Kanpur	133	IC-49586	NBPGR, Jodhpur
40	NBC-51	NBPGR, Delhi	87	AV-6	GKVK, Bengaluru	134	EC-277	IIPR,Kanpur
41	PKB-4-3	GKVK, Bengaluru	88	NBC-21	NBPGR, Delhi	135	EC-458483	IIPR,Kanpur
42	AV-2-1	GKVK, Bengaluru	89	IC-25105	NBPGR, Jodhpur	136	IC-202781	NBPGR, Jodhpur
43	MFC-09-18	VC Farm, Mandya	90	Goa local	IIPR,Kanpur	137	EC-458440	NBPGR, Jodhpur
44	MFC-09-4	VC Farm, Mandya	91	AV-2	GKVK, Bengaluru	138	V-585	IIPR,Kanpur
45	PV-3-2	NBPGR, Jodhpur	92	PKB-4	GKVK, Bengaluru	139	IIHR-133	IIHR, Bengaluru
46	PV-1-2	NBPGR, Jodhpur	93	NBC-32	NBPGR, Delhi	140	EC-458490	IIPR,Kanpur
47	NBC-27	NBPGR, Delhi	94	NBC-30	NBPGR, Delhi	141	NBC-42	NBPGR, Delhi
142	NBC-14	NBPGR, Delhi	152	IC-206240	NBPGR, Jodhpur	162	IC-202781	IIPR,Kanpur

Sl. No.	Genotype	Source	Sl. No.	Genotype	Source	S;. No.	Genotype	Source
143	EC-458453	IIPR,Kanpur	153	EC-170604	IIPR,Kanpur	163	CP-98	IIPR,Kanpur
144	MFC-09-01	VC Farm, Mandya	154	IC-202777	NBPGR, Jodhpur	164	CPD-15	NBPGR, Jodhpur
145	IC-257428	NBPGR, Jodhpur	155	EC-458490	IIPR,Kanpur	165	NBC-12	NBPGR, Delhi
146	EC-458511	NBPGR, Jodhpur	156	C-33	IIPR,Kanpur	166	IC-202825	IIPR,Kanpur
147	IC-249141	NBPGR, Jodhpur	157	IC-402135	NBPGR, Jodhpur	167	IC-402175	IIPR,Kanpur
148	IT-97K499-38	IIPR,Kanpur	158	NBC-21	NBPGR, Delhi	168	EC-458418	IIPR,Kanpur
149	PKB-6	IIPR,Kanpur	159	EC-458483	NBPGR, Jodhpur	169	IC-402159	NBPGR, Jodhpur
150	IT38956-1	IIPR,Kanpur	160	EC-458480	NBPGR, Jodhpur			
151	KM-5	IIPR,Kanpur	161	V-604-7-29-3	NBPGR, Jodhpur			

groups the sets of germplasm accessions, especially valuable for screening large number of descriptor variables. The tenet of present study was to find out the genetic diversity among collected genotypes and to identify better combinations as selection criteria for developing high yielding cowpea genotypes, which may help the plant breeders by promoting new opportunities for development of cowpea cultivars with better yield.

Material and Methods

The experimental material consisted of 169 cowpea accessions which were collected from research stations of Karnataka [MRS (Hebbal), UAHS (Shivamogga), College of Agriculture, VC Farm (Mandya) and AICRP on Arid Legumes, ZARS, UAS (Bengaluru) Table 1].

Experimental design and data analysis

The entries were planted during August 2015 at College of Sericulture Chintamani (weather data given in Supplementary Table 2) and field experiment was laid out in a 13 x 13 simple lattice design with two replications. Each replication consisted of thirteen sub-blocks with thirteen genotypes in each sub block. Entries and sub blocks were randomized. Each genotype was grown in one row of 2-meter length. A spacing of 45 cm between row and 15 cm between plants was maintained. Five randomly selected plants from each the 169 genotypes were used for recording observations on days to first flowering, days to fifty *percent* flowering, plant height, primary branches per plant, clusters per plant, pods per cluster, pods per plant, pod length, seeds per pod, pod yield per plant, 100-seed weight, days to maturity and seed yield per plant. Prior to the initiation of experiment, the surface soil (0-15 cm) sample was collected for the determination of fertility status. Analysis of the sample was carried out in Department of Soil Science and Agricultural Chemistry, UAS, GKVK, Bengaluru. The results are presented in the Supplementary table 1. Dry seeds of cowpea were used to estimate iron content, zinc content and protein content in seeds using Absorption Spectrophotometry (AAS) method and diacid mixture procedure as proposed by Lindsay and Norvell (1978). Five gram of dry cowpea seed flour sample was used analyze crude protein as described in the method no. 920.87 of the Association of Official Analytical Chemists (AOAC, 1995). The cowpea genotypes were classified following model-based k-means clustering approach as suggested by Mac queen, 1967 to unravel organization of variability using SAS 9.3 version software programme. Dendrogram based hierarchical clustering with the help of R function hclust function is constructed. Multivariate approach, D^2 statistic, commonly used for classifying genotypes based on their genetic distance has no control over number clusters to be formed sometimes leads ambiguity in interpretation of results. In case of k-mean clustering, one can customize on number of clusters to be formed thus meaning validation of results are possible. Data ordination multivariate approach, principal component analysis (PCA) computed following Rao (1964) as a linear combination of optimally weighted observed variables.

Results and Discussion

K-means clustering

K-mean clustering is a widely used one of the clustering algorithm helps grouping genotypes into different clusters with control over number of clusters to be formed. One hundred and sixty-nine cowpea genotypes were grouped into eight distinct clusters following ‘k-means’ clustering algorithm (Mac queen, 1967) and are represented in Table 2. Highest number of genotypes were included in cluster VII (56 genotypes) followed by cluster III (43 genotypes), cluster V (35 genotypes), cluster IV (25 genotypes), cluster I (5 genotypes), cluster II (3 genotypes). Clusters VI and

Table 2. List of genotypes under different clusters formed by k-mean clustering from 169 genotypes of cowpea

Cluster	Number of genotypes	Name of genotype
I	5	IC-4506, PV-1-2, MS-4, EC-277, PKB-6
II	3	MFC-09-15, AV-6, IC-25105
III	43	IT38956-1, IC-21483,PKB-4-3, PGCP-3, PGCP-5, Pant lob-1, AV-5, PV-3, PKB-4-2, IIHR-140, PV-1, IC-1061, PKB-5, APC-1218, Goa local, AV-2, IIHR-144, NBC-19, EC-458473, IC-402125, EC-170584, NBC-39, EC-402104, IC-402180, EC-472252, C-24-1, IC-1071, V-16, MFC-09-02, IC-202781, NBC-24, IC-201, IC-402101, NC-32, EC-394839, MBC-25, IC-49586, EC-458483, EC-458440, V-585, NBC-42, EC-458453, IC-249141
IV	25	PV-3-1, MFC-09-16, EC-458505, C-457, MFC-09-8, NBC-44, PV-1-4, IC-402114, Genotype-36, PV-3-4, IC-4506, IC-402161, ETC-27, EC-458473, EC-458489, NBC-51, AV-2-1, MFC-09-18, MFC-09-4, EC-472250, MFC-09-10, EC-170584, MFC-09-05, IC-402180, CP-98
V	35	IC-249593, MFC-09-20, IC-202711(58), NBC-38, C-24-1, NBC-8, 202804(83), PKB-3, NBC-21, C-33, MFC-09-20, IC-202781, IIHR-133, NBC-14, IC-257428, EC-458511, IT-97K499-38, IT-38956-1, KM-5, IC-206240, EC-170604, IC-202777, EC-458490, C-33, IC-402135, NBC-21, EC-458483, EC-458480, V-604-7-29-3, CPD-15, NBC-12, IC-202825, IC-402175, EC-458418, IC-402159
VI	1	PKB-4-1
VII	56	EC-458480, IC-402180, EC-458473, IC-249588, EC-402159, IC-2591054, TOME-774, PV-1-3, AV-2-2, MFC-09-19, AV-5-1, MFC-09-3, V-585, IC-402101, MFC-09-7, PV-3-2, NBC-27, EC-458402, CB-10, PGCP-6, MFC-09-17, PGCP-12, Pant lob-3, Pant lob-2, PGCP-27, PKB-1, PKB-2, IC-201, IIHR-137, MBC-25, NBC-16, IC-402106, EC-458438, Kashi Kanchan, AV-1, IC-58905, PKB-4, NBC-32, NBC-30, Arka Garima, C-720, NBC-47, NBC-43, IC-202781, C-720, CHILORE-11, NBC-30, IC-402090, IT-97154299, IC-202867, EC-170604, 27749(25), GC-3, EC-458490, MFC-09-01, IC-202781
VIII	1	Bhagyalakshmi

Table 3. Mean comparison profile of cowpea germplasm lines into different clusters by K-mean clustering method

Characters	Clusters and number of genotypes in each cluster								F Value	P Value
	C1	C2	C3	C4	C5	C6	C7	C8		
	5	3	43	25	35	1	56	1		
Days to first flowering	58.50	49.00	57.58	48.70	50.93	55.00	50.49	47.50	6.29**	0.00
Days to 50% flowering	61.70	51.83	60.69	51.68	53.90	58.00	53.44	50.50	6.33**	0.00
Plant height (cm)	108.64	75.13	128.13	89.29	81.15	74.15	79.57	112.26	23.95**	0.00
Primary branches/plant	5.46	5.43	5.33	5.43	5.19	5.13	5.11	6.88	1.09	0.37
Clusters/plant	8.45	8.84	7.75	9.71	9.72	6.75	8.12	5.63	1.70	0.11
Pods/cluster	1.56	1.44	1.47	1.61	1.57	1.38	1.49	1.19	2.73*	0.01
Pods/plant	13.29	14.29	12.56	16.64	16.92	9.00	13.29	8.13	2.57*	0.01
Pod length (cm)	16.02	17.11	16.77	16.95	16.60	20.11	16.46	15.43	0.56	0.78
Seeds/pod	13.07	13.24	13.92	13.62	13.80	15.13	13.28	14.63	0.50	0.83
Pod yield/plant (g)	20.61	13.15	17.98	26.34	26.71	11.34	19.90	12.65	3.14**	0.00
100-seed weight (g)	10.80	10.83	10.26	10.70	10.30	12.00	10.69	11.00	0.37	0.91
Days to maturity	89.60	86.83	87.92	87.00	87.04	88.00	85.96	85.00	1.07	0.38
Iron content (ppm)	11.81	7.61	9.78	14.03	14.47	4.96	10.77	7.34	3.05**	0.00
Zinc content (ppm)	384.76	616.53	122.65	218.04	57.69	231.20	133.12	848.00	382.99**	0.00
Protein content (%)	18.00	18.67	35.66	14.51	16.65	192.00	28.10	17.40	34.97**	0.00
Seed yield per plant (g)	23.81	24.35	24.58	24.01	24.10	25.54	24.34	23.40	0.97	0.45

VIII were solitary clusters. The traits mean differences between clusters were significant for most of the traits (Table 3) except for primary branches per plant, clusters per plant, pod length, seeds per pod, hundred seed weight, days to maturity and protein content. The estimates of the means of pod length, seeds per pod, hundred seed weight, zinc content and protein content were highest among the genotypes were included in cluster VI. While, days to first flowering, days to fifty *percent* flowering, clusters per plant, pods per cluster, pods per plant, pod length, days to maturity and protein content had lowest estimates of the means for genotypes grouped under cluster VIII. Genotypes from different geographical origin distributed into all the clusters indicating that there is no relationship between geographical distribution and genotypic diversity. Though wide range of phenotypic variability observed in the study, clustering of genotypes from different sources grouped under many clusters might be due less genetic distance between them.

Inter cluster values and Inter relation of clusters

It would be desirable to attempt crosses between genotypes belonging to distant clusters for getting highly heterotic crosses (Venkatesh et al., 2017). But, heterosis cannot be exploited in a highly self-pollinated crop like cowpea (Yadav et al., 2010) on commercial scale. However, the crosses involving parents from cluster with high inter cluster distance are likely to yield desirable recombinants in the advanced generation, which could be developed as traditional homozygous varieties. In this context, inter cluster distances were worked out considering the quantitative characters and the distances varied from 51.35 (between clusters III and VII) to 791.50 (between V and VIII). The inter cluster distances was maximum between cluster V and VIII (791.15) followed by cluster III and VIII (725.94) and between cluster VII and VIII (715.79). These values suggest that the genotypes from distant clusters exhibit wide diversity (Falconer, 1981). Thus, genotypes from divergent clusters can be selected for hybridization programme for obtaining desirable recombinants.

The nearest and farthest cluster from each individual clusters are presented in Supplementary table 3. The inter-cluster distances varied from 51.35 (between cluster III and VII) to 791.15 (between clusters V and VIII). All the other inter-cluster distances were lying between these values. Cluster I consisting of 5 genotypes was nearer to clusters IV (168.63). Cluster II comprised of 3 genotypes was close to cluster VIII (234.56). It was farthest from cluster V (559.10). Cluster III, IV and V were grouped with 43, 25 and 35 genotypes respectively, which were closer to clusters VII (51.35, 86.97 and 76.81). Cluster VI had only one genotype in its group. It was nearer to clusters IV (179.93). Clusters VII comprised of 56 genotypes and was nearer to cluster III (51.35). Cluster I, III, IV, V, VI and VII are farthest from cluster VIII (463.67, 725.94, 630.65, 791.15, 642.29 and 715.79 respectively). Cluster VIII was grouped with single genotype, which was closer to the cluster II (234.56) and was farthest from cluster V (791.15).

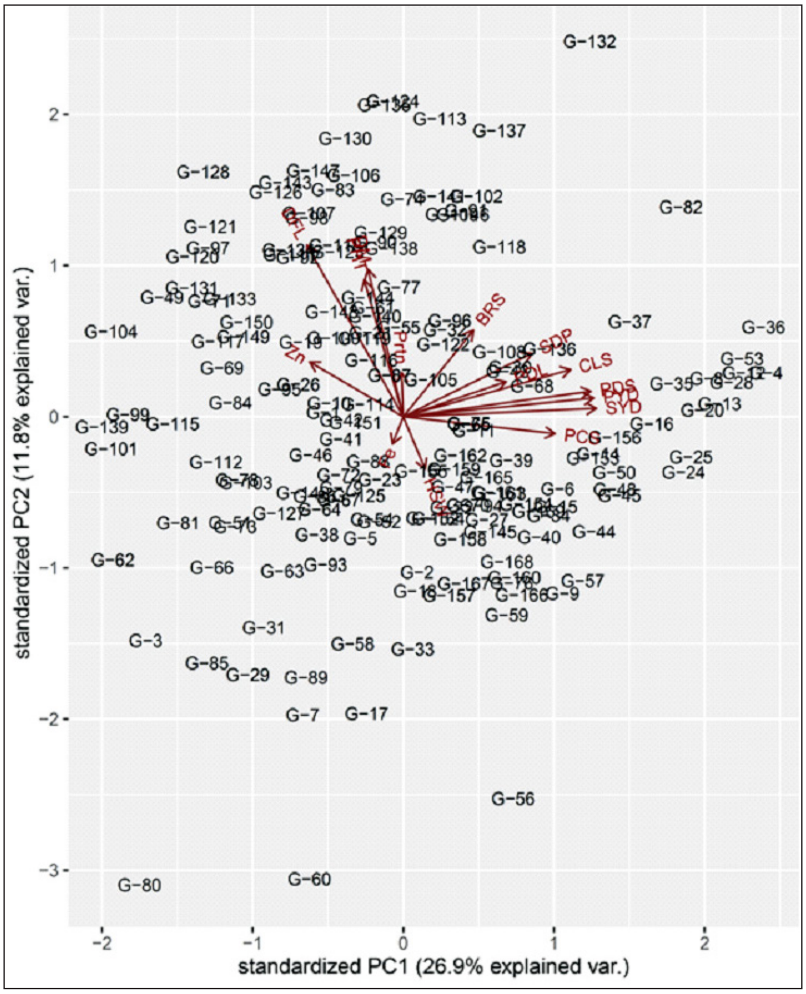


Fig. 2. Distribution of genotypes across PC1, PC2

Cluster mean analysis

Cluster mean values for all characters are presented in Table 3 indicate that clusters differ with respect to mean performance of zinc, iron, protein, yield and its attributes. This reflects the clusters formed are very distinct. The genotypes grouped under cluster I and II consisting of five and three genotypes respectively showed high mean values for days to first flowering, days to *percent* flowering and days to maturity. Cluster VI with single genotype (PKB-4-1) represented high mean values compared to other clusters for character pod length, seeds per pod, hundred-seed weight, zinc content and protein content. Two clusters (VI and VIII) with single genotypes each were widely diverse for days to first flowering, plant height, pod length, seed yield per plant, iron content and zinc content. Genotype Bhagyalakshmi recorded lowest cluster mean for days to first flowering, days to fifty *percent* flowering, clusters per plant, pods per cluster, pods per plant, pod length, days to maturity and protein content was grouped under cluster VIII.

Highest mean value for pods per cluster was recorded for genotypes classified under cluster IV. It is worthy to note that calculating cluster means, the superiority of a particular genotype in respect of a given character get diluted by other genotype that are related and grouped in the same cluster which are interior or intermediary for that character in question. Hence, apart from selecting lines from cluster which have high cluster distance for hybridization, one can also think of selecting parents based on the extent of divergence in respect to a character of interest. This means that if breeder's intention is to improve the seed yield to plant then selection should be aimed at parents which are highly divergent for the trait of interest. The clusters VI and VIII that are having diverse mean values of most of the characters can be used as parents in hybridization programme aimed at improving the trait in consideration. These findings were in line with the reports of Dalsaniya *et al*. (2009) and Sandeep *et al*. (2014).

Principal component analysis

Principal component analysis was performed to identify the most contributing variables to the total variability observed (Fig. 2). Five principal components with Eigen values greater than one contributed 70.5 *percent* of the total variability among 169 genotypes evaluated for 16 traits including three nutritional traits. The following traits contributed positively to PC_1 that had 31.3 *percent* variation proportion to the total variability: pod yield per plant (0.408), pods per plant (0.401) and seed yield per plant (0.396). In PC_2 accounting for 14.6 *percent* of the total variability, days to first flowering (0.512), days to fifty *percent* flowering (0.511) and days to maturity (0.430) contributed positively to the total variation.

The traits pod length (0.581), seeds per pod (0.487) and zinc content (0.341) contributed positively to the total variation, which is noticed in PC_3 that accounts for 10.10 *percent* of total variation. PC_4 had 7.7 *percent* variation of the total variation where iron content (0.583) and 100-seed weight (0.452) contributed positively and negatively by protein content (-0.463) to the variation. In PC_5, protein content (0.563) and primary branches per plant (0.484) contributed positively while 100-seed weight (-0.359) contributed negatively to the variation. From the entire five principal components pod yield per plant, seed yield per plant, pod length and seeds per pod contributed significantly to the total genetic variability and divergence.

It revealed that pod yield per plant, seed yield per plant, pod length and seeds per pod contributed significantly to the total genetic variability and divergence. These results corroborate those of earlier reports by Singh *et al.* (2008), Muhammad lawan umar (2014) and Udensi and Edu (2015).

Conclusions

K-means clustering was opted to organize the genetic variability present in a set of 169 genotypes. The values ranged from 51.35 to 791.15. Grouping the genotypes into clusters using K-means clustering method resulted in the formation of 8 clusters of which cluster VII was the biggest with 56 entries followed by cluster III, V, IV and I with 43, 35, 25 and 5 genotypes respectively. The mean differences between clusters were significant for the traits days to first flowering, days to fifty *percent* flowering, plant height, pods per cluster, pods per plant, pod yield per plant, seed yield per plant, iron content and zinc contents. The inter cluster distance varied from 51.35 (between clusters III and VII) to 791.15 (between V and VIII). The estimates of the means of the traits such as pod length, seeds per pod, hundred seed weight, zinc content and protein content were highest among the genotypes included in cluster VI. Traits like days to first flowering, days to fifty *percent* flowering, clusters per plant, pods per cluster, pods per plant, pod length, days to maturity and protein content had lowest estimates of the means in cluster VIII. Thus genotypes from the clusters VI and VIII are found to be genetically divergent and contrasting which could be utilized in the hybridization programme.

Principal component analysis identified the most contributing variables to the total variability. Five principal components with Eigen values greater than one contributed 70.5 *percent* of the total variability among 169 genotypes evaluated for 16 traits including three nutritional traits. Pod yield per plant, seed yield per plant, pod length and seeds per pod contributed significantly to the total genetic variability and divergence. Thus, priority should be given these characters when selection is exercised in cowpea crop improvement programmes. Performing genetic divergence studies before starting up a new breeding programme will help to determine the amount of genetic diversity present in it, thereby reducing the cost of money and time in crop improvement programmes.

Estimates of iron, zinc and protein aided in identification of nutrient efficient genotypes in the selected population of 169 cowpea genotypes. Based on the mean composition of these nutrient traits, genotypes were classified into different categories. Twelve genotypes (PKB-4-1, MFC-09-17, PKBB-2, EC-170584, NBC-47, C-720, CHILORE-11, NBC-39, EC-402104, MFC-09-05, 27749-25 and NBC-24) identified as zinc, iron and protein rich.

Acknowledgement

MRS (Hebbal), UAHS (Shivamogga), College of Agriculture, VC Farm (Mandya) and AICRP on Arid Legumes, ZARS, UAS (Bengaluru).

*Supplementary Tables or Figures mentioned in the article are available in the online version.

References

Association of Official Analytical Chemists AOAC (1995) Official method 920.87 – Protein (total) in flour, final action. In: P. Cunniff, editor, Official methods of analysis of the association of official analytical chemists. 16th ed. AOAC, Arlington, VA. Vol. 2, chapter 32, p. 12-13.

Dalsaniya SB, VK Poshiya, JJ Savaliya, AG Pansuriya and BK Davada (2009) Genetic divergence in cowpea [*Vigna unguiculata* (L.) Walp.]. *Legume Res.* **32(4):** 250-254.

Falconer DS (1981) Introduction to Quantitative Genetics, Ed. 2. Longmans Green, London/New York.

Fisher RA (1981) The correlation among relative on the supposition of Mendelian Inheritance. Transactions of the Royal Society of Edinburgh **52:** 399-433.

Howarth E. Bouis, Ross M. Welch (2010) Biofortification-A Sustainable Agricultural Strategy for Reducing Micronutrient Malnutrition in the Global South. Crop Sci. **50** :S20-S32. https://doi.org/10.2135/cropsci2009.09.0531

https://www.nutritionvalue.org/Cowpeas%2C_raw%2C_mature_seeds%2C_catjang_nutritional_value.html?size=100+g

Ira A Herniter, María Muñoz-Amatriaín, Timothy J Close (2020) Genetic, textual, and archeological evidence of the historical global spread of cowpea (*Vigna unguiculata* [L.] Walp.), Legume Sci. **2**:e57. https://onlinelibrary.wiley.com/doi/10.1002/leg3.57

Lindsay WL, Norvell WA (1978) Development of a DTPA soil test for zinc, iron, manganese, and copper. *Soil Science Society of America J.* **42**: 421-428.

Mac Queen JB (1967) Some Methods for classification and analysis of multivariate observations, Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, University of California Press **1:** 281-297.

Muhammad Lawan Umar (2014) Breeding for grain quality traits in cowpea [*Vigna unguiculata* (L) walp]. Ph D thesis, University of Ghana, Legon.

Rao CR (1964) The use and interpretation of principal component analysis in applied research. Sankhya **26:** 329-358.

Sandeep VV, Hemalatha BD Shashi and V Swarnalatha (2014) Studies on genetic diversity in Indian cowpea (*Vigna unguiculata* (L.) Walp) germplasm. *Int. J. Plant Animal Environ. Sci.* **4(3):** 177-180.

Singh SK, GK Surabhi, W Gao and KR Reddy (2008) Assessing genotypic variability of cowpea (*Vigna unguiculata* (L.) Walp.) to current and projected ultraviolet-B radiation. *J. Photoch Photobio*. **93:** 71-81.

Udensi OU and NE Edu (2015) Evaluation and identification of genetic variation pattern in cowpea [*Vigna unguiculata* (L.) Walp] accessions using multivariate analyses. *J. Basic Appl. Sci.* **11:** 149-158.

Venkatesh L, Niranjana Murthy and SD Nehru, (2014) Analysis of genetic diversity in grain amaranth (*Amaranthus* spp.). *Indian J. Genet*. **74 (4)**: 522-525.

Yadav KS, Yadava HS and Dixit H (2010) Heterosis and inbreeding depression in cowpea. *Intl.J. Agril. Sci.***6 (2)**: 537-540.

Supplementary Table 1. Physico-chemical properties of soil of the experimental site

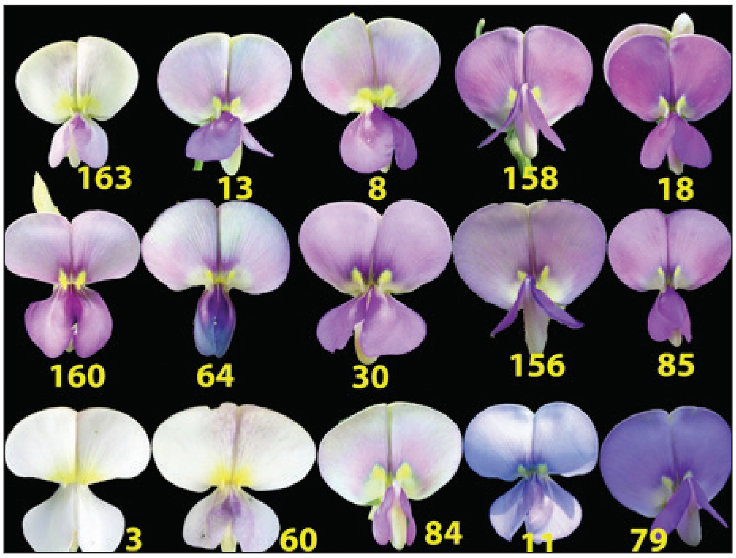
Sl. No.	Physico-chemical properties	Results	Sl. No.	Physico-chemical properties	Results
I.	**Physical properties**		**II.**	**Chemical properties**	
1	Coarse sand (%)	48.15	1	pH	6.35
2	Fine sand (%)	28.30	2	Electrical conductivity (dSm^{-1})	0.32
3	Silt (%)	14.15	3	Organic carbon (%)	0.61
4	Clay (%)	9.40	4	Available N (kg ha^{-1})	368
5	Soil textural class	Sandy clay loam	5	Available P_2O_5 (kg ha^{-1})	41.3
6	Field capacity (%)	16.60	6	Available K_2O (kg ha^{-1})	298
7	Permanent wilting point (%)	5.80	7	Iron (ppm)	26.7
8	Bulk density (g cc^{-1})	1.35	8	Zinc (ppm)	2.5

Supplementary Table 2. Meteorological data of Agriculture Research Station, Chintamani from January to December 2015

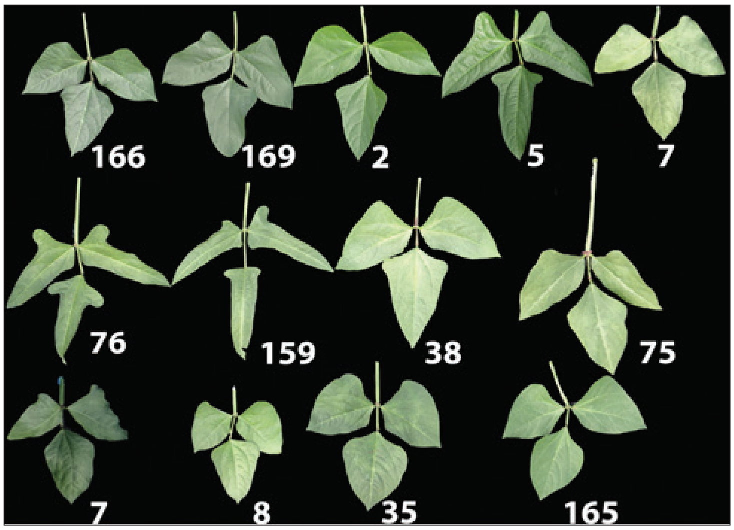
Month	Total Rain Fall (mm)	No. of rainy days	Temperature (°C)		Relative Humidity (%)		Water Evp	SS Hours
			Minimum	Maximum	Morning	Evening		
January	6.2	1	16.86	27.66	73.00	39.12	3.49	4.54
February	–	–	17.29	29.99	63.70	34.28	5.30	9.66
March	35.2	4	21.32	32.93	68.77	49.48	5.27	8.28
April	170.5	5	21.95	33.39	64.70	48.96	4.60	7.94
May	160.9	6	23.12	31.78	65.80	51.25	3.45	6.95
June	40.1	5	21.47	34.95	70.60	57.16	3.05	5.68
July	60.0	5	20.53	28.44	65.12	53.96	2.98	6.48
August	162.3	10	20.61	30.35	72.83	59.48	2.01	6.10
September	189.8	8	20.27	30.08	71.20	56.83	2.47	5.99
October	113.3	7	19.08	30.29	71.54	59.54	2.76	7.69
November	349.1	17	17.10	26.92	79.86	72.23	1.51	3.26
December	3.0	–	16.03	27.80	74.58	64.90	3.27	6.57
Total	1284.2	68						

Supplementary Table 3. Principal component analysis showing the contribution of morphological and nutritive traits to the total variation among the cowpea accessions

Communality	PC_1	PC_2	PC_3	PC_4	PC_5
Days to first flowering	-0.243	0.512	-0.062	-0.003	-0.079
Days to 50% flowering	-0.244	0.511	-0.063	-0.007	-0.080
Plant height	-0.105	0.301	0.263	0.016	-0.081
Primary branches per plant	0.141	0.221	-0.022	0.351	0.484
Clusters per plant	0.376	0.121	-0.051	-0.030	0.149
Pods per cluster	0.289	0.067	-0.252	-0.170	-0.131
Pods per plant	0.401	0.134	-0.155	-0.094	0.086
Pod length	0.192	0.074	0.581	0.056	-0.098
Seeds per pod	0.230	0.175	0.487	-0.054	-0.015
Pod yield per plant	0.408	0.168	-0.050	-0.007	-0.084
100-seeds weight	0.068	-0.065	0.226	0.452	-0.359
Days to maturity	-0.089	0.430	-0.228	0.179	0.006
Seed yield per plant	0.396	0.164	-0.041	0.004	-0.126
Iron content	-0.014	-0.091	0.040	0.583	0.458
Zinc content	-0.187	0.082	0.341	-0.198	0.095
Protein content	-0.020	0.056	0.172	-0.463	0.563
Eigen value	5.009	2.337	1.608	1.234	1.086
Proportion of variation (%)	31.3	14.6	10.1	7.7	6.8
Cumulative variance (%)	31.3	45.9	56.0	63.7	70.5



(a) Variability Observed for flower colour



(b) Variability Observed for leaf shape

Supplementary Fig. 1. Variability Observed for flower colour (a); (b) Variability Observed for leaf shape