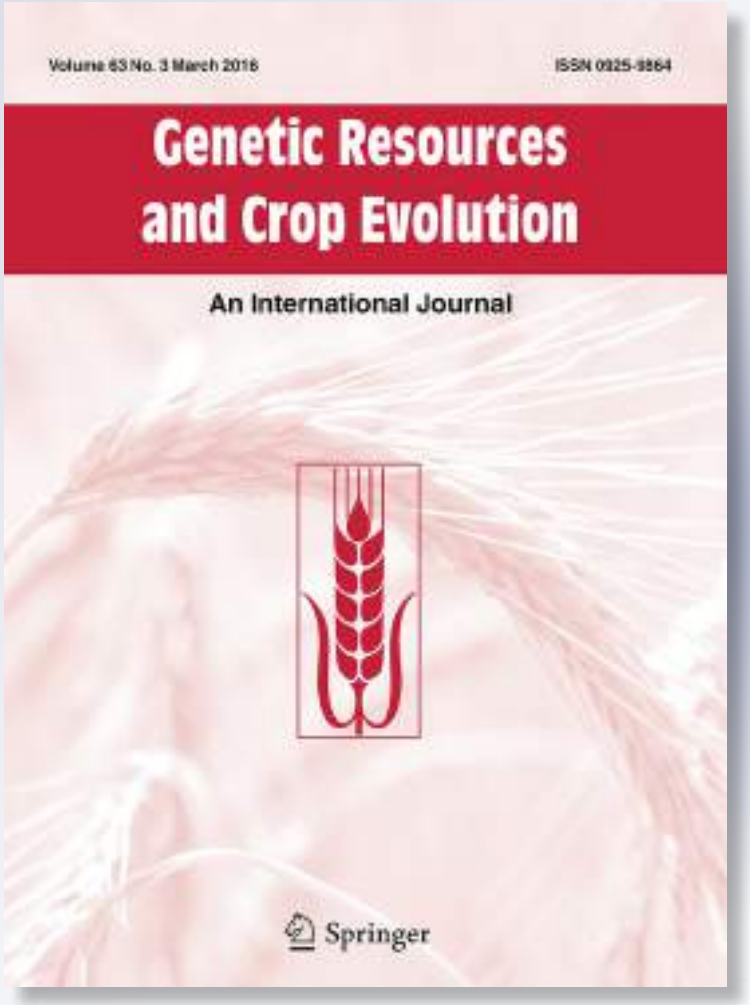
*Characterization of Indian clusterbean (*Cyamopsis tetragonoloba *(L.) Taub.) genotypes using qualitative morphological traits*

A. Manivannan, C. R. Anandakumar, R. Ushakumari & G. S. Dahiya

Genetic Resources and Crop Evolution
An International Journal

ISSN 0925-9864
Volume 63
Number 3

Genet Resour Crop Evol (2016)
63:483-493
DOI 10.1007/s10722-015-0266-y



Springer

Genet Resour Crop Evol (2016) 63:483–493
DOI 10.1007/s10722-015-0266-y

RESEARCH ARTICLE

Characterization of Indian clusterbean (*Cyamopsis tetragonoloba* (L.) Taub.) genotypes using qualitative morphological traits

A. Manivannan · C. R. Anandakumar · R. Ushakumari · G. S. Dahiya

Received: 11 February 2015 / Accepted: 5 May 2015 / Published online: 22 May 2015

Abstract Clusterbean (*Cyamopsis tetragonoloba* (L.) Taub.) is commonly known as *guar*, a legume with drought hardiness suited to arid regions remained a neglected and underutilized crop till the shale rush began, of late it got momentum in the international market, and it became one of the most promising industrial crops in India. Seeds of clusterbean are used for extraction of gum which is called *guar* gum or guaran. *Guar* gum—a galactomannan polysaccharide—has many uses such as in the oil and petroleum industries (shale energy production or oil fracking), food industries, pharmaceutical industries, paper industries and mining fields. India is the major producer and exporter of guar to the world market. There is a tremendous variability available in Indian clusterbean germplasm, our objective of the present study to characterize and evaluate the Indian clusterbean genotypes using morphological qualitative traits. Forty-two genotypes collected across India were characterized for seven morphological traits: stem type, growth habit, leaflet texture, leaflet margin, leaflet size, flower colour and seed colour across three locations during *kharif* (rainy season) 2013. Coefficients of variation and principal component analysis revealed variability among the genotypes for the qualitative traits evaluated. PCA analysis showed that genotypes namely PNB, T local, HGS 884, RGC 471, MRSG6 were very different from each other. Cluster analysis separated genotypes into two major groups: one of vegetable genotypes (PNB, T local, M local, HVG2-30 and Amrit 11), and the other of gum genotypes. Among the gum genotypes RGC1066, RGC197 and FS277 formed a distinct sub cluster as they possess single stems (unbranched type), while the genotypes with white flower colour namely RGC936 and RGC471 formed a different sub cluster. The genotype MRSG6 showed brown seed with distinct grade of N 200 B which was distinctly unique and differed from the predominant grey-seeded group Correlation analysis among the traits showed that leaflet texture (pubescence) and leaflet shape (narrow) correlated highly with gum genotypes, on the contrary glabrous broad leaflets were linked with vegetable genotypes.

A. Manivannan (✉)
Indian Council of Agricultural Research - Indian Institute of Maize Research, Pusa, New Delhi 110012, India
e-mail: manivannan461@gmail.com

C. R. Anandakumar
Centre for Plant Breeding and Genetics, Tamil Nadu Agricultural University, Coimbatore 641003, India

R. Ushakumari
Department of Plant Breeding and Genetics, Agricultural College and Research Institute, Tamil Nadu Agricultural University, Madurai 625104, India

G. S. Dahiya
Department of Genetics and Plant Breeding, College of Agriculture, Chaudhary Charan Singh Haryana Agricultural University, Hisar 1250004, India

Keywords Clusterbean · *Guar* gum · Morphological traits · Cluster analysis · Principal component analysis · Variability · Dendrogram

Introduction

Clusterbean (*Cyamopsis tetragonoloba* (L.) Taub.) is commonly known as *Guar*. It is a deep rooted annual arid legume crop highly valued for its industrial applications in the market. It is known for its drought tolerance and highly adopted to various moisture-limited environments. Clusterbean originated from the Indian sub-continent of Asia and is grown over arid and semiarid regions of the world (Purseglove 1981). India is the major producer accounting for 80 % of the world's production followed by Pakistan, USA and South Africa. In India, clusterbean is being cultivated mainly in arid and semiarid regions covering about 5.15 million hectares with a production of 2.46 million tonnes with productivity of 478 kg/ha in 2012–13 (Department of Agriculture and Co-operation, MoA, GoI 2014; NAIM 2014). India is the major exporter of *guar* gum to the world; it exports various forms of *guar* products *such* as *guar* gum, *guar* split, *guar* meal to a large number of countries. The exports of *guar* holds a share of 18 % among the agricultural commodities during 2012–13 (NIAM 2014). The country has exported 0.41 million tonnes of *guar* gum to the world for the worth of 3.85 billion US$ during the year 2012–13 (APEDA 2014). Major export destinations are United States, China, Germany, Canada, Russia and Australia (APEDA 2014).

Seed of clusterbean has rather a large endosperm unlike most of other legume seeds and contains gum which is a galactomannan polysaccharide. Galactomannan is a natural thickener that is water-soluble at low temperatures (hydro-colloidal in nature). Because of its viscous property it has emerged as an important industrial gum which has many applications, viz, Oil and Gas well drilling (shale industries), Paper and Textile industries, Food industries, Cosmetics, Mining, Pharmaceuticals, etc. Clusterbean gum or *guar* gum has emerged as the most important agro-chemical, which is non-toxic, eco-friendly and Generally Recognized As Safe (GRAS) by FDA.

Variation in germplasm collections has been utilized for identifying desirable genotypes to enhance yield improvement. To identify the desired genotype, various morphological traits are being employed such as leaflet type, flower colour, growth habit, leaflet texture, leaflet shape, canopy pattern, seed size, etc. (Emami and Sharma 1999; Tanksley 1983). These distinct qualitative traits are called morphological markers, often reliable for germplasm characterization and trait associated selections; such qualitative traits are stable in expression across the environment. Characterization of germplasm plays a vital role in crop improvement.

Characterization, evaluation of germplasm and quantification of genetic diversity genotypes is indispensable for a pragmatic use of plant genetic resources and also for determining evolutionary relationships (Zada et al. 2013). Studies of the variation present in germplasm collections have been carried out employing plant morphological attributes as characterization tool among Leguminosae pulses such as Butterfly pea (*Clitoria ternatea*) (Morris 2009), Fenugreek (*Trigonella foenum-graecum*) (McCormick et al. 2009), Sword-bean (*Canavalia ensiformis)* (Morris 2007), White clover (*Trifolium repens*) (Rosso and Pagano 2001), Alfa alfa (*Medicago sativa*) (Rumbaugh et al. 1988), White lupin (*Lupinus albus*) (Rubio et al. 2004) and Horsegram (*Macrotyloma axillare*) (Morris 2008).

The present study was envisaged to characterize clusterbean genotypes collected across the Indian sub-continent for qualitative traits namely stem type, growth habit, leaflet surface, leaflet margin, leaflet shape, flower colour and seed colour using correlation analysis, principal component analysis, cluster analysis and diversity analysis.

Materials and methods

Genetic evaluation of clusterbean genotypes on the basis of agro-morphological traits was performed at three locations in the state of Tamil Nadu: the Agricultural College and Research Institute, Madurai; the Centre for Plant Breeding and Genetics, Tamil Nadu Agricultural University, Coimbatore and the Agricultural Research Station, Kovilpatti. Germplasm accessions (Table 1) were obtained from various leading centers of clusterbean research which represent eco-geographical diverse areas of India. Experiment was conducted as per RCBD (Randomized Complete Block Design) with plot sizes of 3×3 m^2 during *Kharif* season, 2013 by two replications per

Table 1 List of genotypes used in the study

S. No	Genotypes	Origin
1	RGC 197, RGC 471, RGC 936, RGC 986, RGC 1002, RGC 1003, RGC 1017, RGC 1031, RGC 1033, RGC 1038, RGC 1055, RGC 1066, RGM1 and RGM 2	Rajasthan Agricultural Research Institute, Durgapur, Rajasthan
2	HGS16, HGS 75, HGS 182, HGS 258,HGS 365, HGS 563, HGS 832, HGS 870, HGS 884, HGS 2-1, HGS 2-4, HGS 2-20, HGS 3-2, HGS 3-52, HFG 119, HVG 2-30 and FS 277	Chaudhary Charan Singh Haryana Agricultural University, Hisar, Haryana
3	CAZG 10-2, MRG 1786, MRSG 6 and SRG 1058	Central Arid Zone Research Institute, Jodhpur, Rajasthan
4	GAU 512 and GAU 513	Sardarkrushinagar Dantiwada Agricultural University, Krushinagar, Gujarat
5	PNB	Indian Agricultural Research Institute, Pusa, New Delhi
6	R local	Landrace from Rajasthan
7	M local and T local	Landrace from Tamil Nadu
8	Amrit 11	Local cultivar from Gujarat

entry. Each entry occupied a three-meter row with spacing of 45 × 15 cm. All agronomical practices were followed as prescribed by the agronomists.

Qualitative characters were taken under consideration for evaluating the substantial variation and relationship among clusterbean genotypes. Seven characters: stem type, growth habit, leaflet surface, leaflet type, leaflet margin, flower colour and seed colour were used to describe genetic diversity. Scores were assigned for each trait as per the guide lines by USDA, ARS, National Genetic Resources Program, Germplasm Resources Information Network (GRIN). Stem type was based on a rating scale of 1–3, where 1 = branched, 2 = not branched and 3 = mix of branched and unbranched (Fig. 1A). Crop canopy stand behaviour was considered as growth habit, the growth was observed for bushy (score = 1) and erect canopy pattern (score = 3). Leaflet surface was based on a rating scale of 1–3 where 1 = glabrous and 3 = pubescent types. The margin of leaflet was observed for the presence of serrated margin (score = 1) and smooth margin pattern (score = 3) (Fig. 1D). The size and shape of the leaflet was recorded as narrow shape (score = 1) and broad (score = 3) (Fig. 1C). Colour of the flower was recorded by using a Royal Horticultural Society (RHS) colour chart; purple colour was given score of 1 and white as 3 (Fig. 1B). Seed colour was categorized based on a Royal Horticulture Society (RHS) colour chart, Five grades of colour were observed namely grey group (score = 1), grey–brown group (score = 3), greyed–green group (score = 5), greyed–yellow group (score = 7) and brown group (score = 9). Since these traits are qualitative in nature, pooled scores from all the locations were taken for analysis (Table 2).

Pearson correlation coefficient was worked out for seven qualitative traits and correlation matrix was prepared for comparing different traits. Principal component analyses (PCA) based on seven qualitative traits was performed to find out the relative importance of different traits in capturing the genetic variation in clusterbean. The factors of these traits were used to determine the contribution of each factor towards variation. The standardized values were used to perform PCA using PAST 3 (Hammer et al. 2001). A scree plot was drawn from the eigenvalues associated with a component or factor in descending order versus the number of the component or factor. Scree plot used for visually assess which components or factors explain most of the variability in the data. Shannon–Weaver diversity index (H′) (Shannon and Weaver 1949) was used as a measure of phenotypic diversity and worked out for each trait. A low H′ indicates extremely unbalance frequency classes for an individual traits and a lack of genetic diversity. The index was estimated based on seven qualitative traits using PAST 3.

Dissimilarity matrix based on EUCLIDEAN distance was calculated using these traits by DARwin 5. Most dissimilar and least dissimilar accessions were identified in clusterbean genotype based on dissimilarity matrix. A hierarchical cluster analysis for pooled data was performed using scores of dissimilarity matrix (Ward 1963).

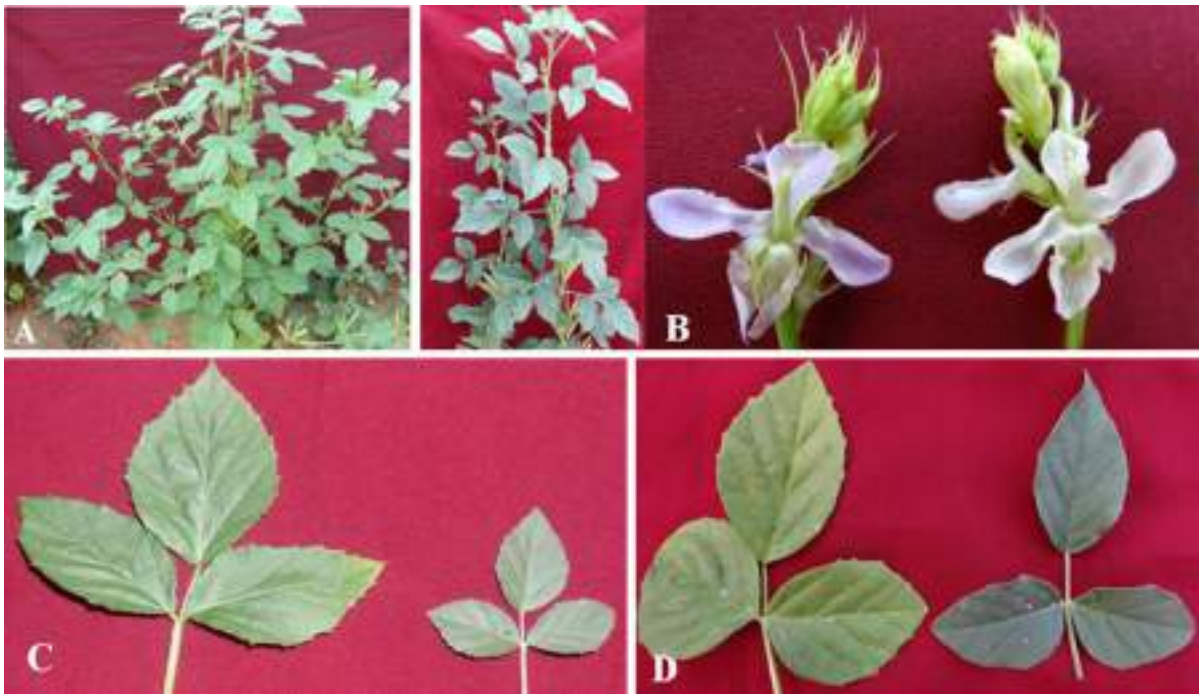


Fig. 1 Morphological trait variation. **A** Stem type (branched and unbranched). **B** Flower colour (purple and white). **C** Leaflet size (broad and narrow). **D** Leaflet margin (serrated and smooth). (Color figure online)

Table 2 Distribution of phenotypic classes among qualitative traits

Descriptor	Score	Category	Number of genotypes	Frequency (%)
Seed colour	1	Grey group	4	9.5
	3	Grey–brown group	29	69
	5	Greyed–green group	4	9.5
	7	Greyed–yellow group	4	9.5
	9	Brown group	1	2.5
Flower colour	1	Purple	39	92.85
	3	White	3	7.15
Leaflet surface	1	Glabrous	37	88.1
	3	Pubescence	5	11.9
Stem type	1	Branched	37	88.1
	3	Unbranched	5	11.9
Leaflet margin	1	Serrate	34	80.95
	3	Smooth	8	19.05
Growth habit	1	Bushy	35	83.34
	3	Erect	7	16.66
Leaflet type	1	Narrow	36	85.7
	3	Broad	6	14.3

Results

Evaluation of genotypes based on qualitative traits

Observations on seven qualitative traits were recorded for all the 42 genotypes and the scores were analyzed (Table 2). Maximum number of genotypes (37) was observed with branched stem type, while five genotypes namely RGC1066, RGC197, FS277, T local and PNB were unbranched (single stem). Two groups namely bushy and erect were observed. Bushy type was found in 35 genotypes and seven genotypes were erect in nature. In the case of seed colour, a maximum number of genotypes (37) were found with pubescent leaflet surface and five genotypes, namely HVG2-30, T local, M local, PNB and Amrit 11, were glabrous. Only two categories of leaflet margin (smooth and serrate) were observed. Among the genotypes, 38 were found to have smooth leaflet margins and eight had serrated margins. Two classes of leaflet type namely narrow and broad were observed. Narrow leaflet types were found to be higher (36) when compared to broad leaflet type (6). Among the genotypes, two categories namely purple and white flowers were observed. Purple colour was found in the majority of genotypes (38). Four genotypes namely HGS16, RGC936, RGC471 and HVG2-30 possess white flowers. Most genotypes possessed grey–brown group (26), however, four genotypes were observed in each of the groups namely grey group, greyed–green group and grey–yellow group. Among them only one genotype, MRSG6 belonged to brown group N 200 B.

Correlation among qualitative traits

Pearson (1901) correlation coefficient was employed among the qualitative traits. Among the 21 inter correlation coefficients, six were significant. The highest significant positive correlation was observed between leaflet type and leaflet surface (0.90). Significant correlation was observed between growth habit and leaflet surface (0.83), leaflet surface and growth habit (0.82), growth habit and leaflet type (0.73), stem type and growth habit (0.63) and leaflet surface and stem type (0.32). All other inter correlations were found to be non-significant (Table 3).

Principal component analysis (PCA)

Four significant principal components were identified and accounted for a cumulative variation of 87.79 %. The first principal component accounted for 43.27 %, second for 17.88 %, third for 14.62 % and fourth for 12.02 % of total variation (Table 4).

Table 3 Correlation among traits

Characters	SC	FC	LS	ST	LM	GH	LT
SC	1.00	0.12	−0.22	−0.09	−0.22	−0.22	−0.22
FC		1.00	−0.1	−0.1	0.1	−0.12	−0.11
LS			1.00	0.32[a]	−0.18	0.82[b]	0.90[b]
ST				1.00	−0.18	0.63[b]	0.27
LM					1.00	−0.22	−0.02
GH						1.00	0.73[b]
LT							1.00

SC seed colour, *FC* flower colour, *LS* leaflet surface, *ST* stem type, *LM* leaflet margin, *GH* growth habit, *LT* leaflet type

[a] Significance at 5 % level

[b] Significance at 1 % level

Table 4 Principal Component Analysis of different traits

Principal component	Eigen value	Proportion of variation (%)	Cumulative
1	3.03	43.27	43.27
2	1.25	17.88	61.15
3	1.09	14.62	75.77
4	1.00	12.02	87.79
5	0.63	9.06	96.85
6	0.15	2.17	99.02
7	0.07	0.98	100.00

Contribution of character towards principal component factors

The scores of qualitative traits were taken into account and subjected to PCA using PAST 3. Eigenvectors and principal components based on non-rotated loadings were estimated for all the qualitative traits (Table 5). First principal component (PC 1) was correlated with growth habit (0.94), leaflet surface (0.92), leaflet type (0.87) and stem type (0.60). The PC 2 was associated with seed colour (0.71), stem type (0.23) and growth habit (0.08). The PC 3 was correlated with flower colour (0.90), seed colour (0.28) and leaflet type (0.21). The PC 4 was associated with flower colour (0.70), stem type (0.30) and growth habit (0.12).

Based on all the principal components, maximum variation was recorded for growth habit (43.82 %), moderate variability (33.43–39.06 %) was recorded for leaflet surface, stem type and leaflet type, while flower colour and seed colour was found to be less in variability (Table 6). Based on the four PCs, scatter and scree plots depicted the variability.

Table 5 Principal components of various traits

Characters	Principal components			
	PC1	PC2	PC3	PC4
Seed colour	−0.30	0.71	0.28	−0.24
Flower colour	−0.21	0.01	0.90	0.30
Leaflet surface	0.92	−0.03	0.19	−0.27
Stem type	0.60	0.23	−0.15	0.70
Leaflet margin	−0.22	−0.81	0.17	0.07
Growth habit	0.94	0.08	0.06	0.12
Leaflet type	0.87	−0.16	0.21	−0.32

Table 6 Qualitative traits contribution towards variability

Characters	Variability contribution (%)
Growth habit	43.82
Leaflet surface	39.06
Stem type	34.48
Leaflet type	33.43
Flower colour	2.03
Seed colour	1.20

Grouping genotypes based on PCA biplot

A scatter plot (Fig. 2) drawn using PC1 and PC2 factor scores (Table 5) and clear pattern of grouping between the genotypes was observed in the factor plane. Convex of the hull occupied by the genotypes namely PNB, T local, HGS 884, RGC 471 and MRSG6; these genotypes showed the highest point among the factors. The genotype Amrit 11 placed in proximity to HVG 2-30, similarly the genotype FS 277 placed in proximity to RGC 197. The genotypes namely RGM1, RGC1003 and RGC 1017 placed in a single point, similarly the genotypes namely HGS 2-1, HGS 3-2 and HGS 832 also grouped together in a single position. Sixteen genotypes namely RGC1002, GAU512, HGS365, HGS75, HGS3-52, RGM2, SRG1058, HGS2-20, RGC1055, CAZG10-2, MRG1786, RGC1038, R local, GAU513, RGC1031, and RGC1033 congregated in a single position.

Cluster analysis

The factors corresponding to four PCs were subjected to cluster analysis based on Euclidean distances and grouped by unweighted paired group method using arithmetic average (UPGMA) using DARwin 5. The dendrogramme depicted two distinct clusters. The cluster I showed two groups consisted of five genotypes, while cluster II showed four groups comprised of 37 genotypes. Group 1, 2, 3 and 4 of cluster II consisted of 3, 3, 1 and 30 genotypes respectively (Table 7; Fig. 3).

Shannon–Weaver diversity index (H′)

The H′ of seven qualitative traits was used for estimating the frequency distribution of diversity index (Table 8). The H′ of the trait flower colour was observed to be higher among (3.67) other traits. An average index (H′) of 3.63 was observed among all the traits.

Phenotypic dissimilarity matrix

Higher level of dissimilarity was found between the pairs of genotype namely T local and RGC936 (7.78), T local and RGC471 (7.49), PNB and RGC471 (7.49), T local and MRSG6 (7.44), RGC 936 and M local (7.16), RGC 936 and PNB (7.05), PNB and HGS16 (7.05), T local and HGS16 (7.05), RGC 471 and M local (6.84), M local and MRSG6 (6.79). However, 81 pairs of genotype were found to possess zero dissimilarity indices.

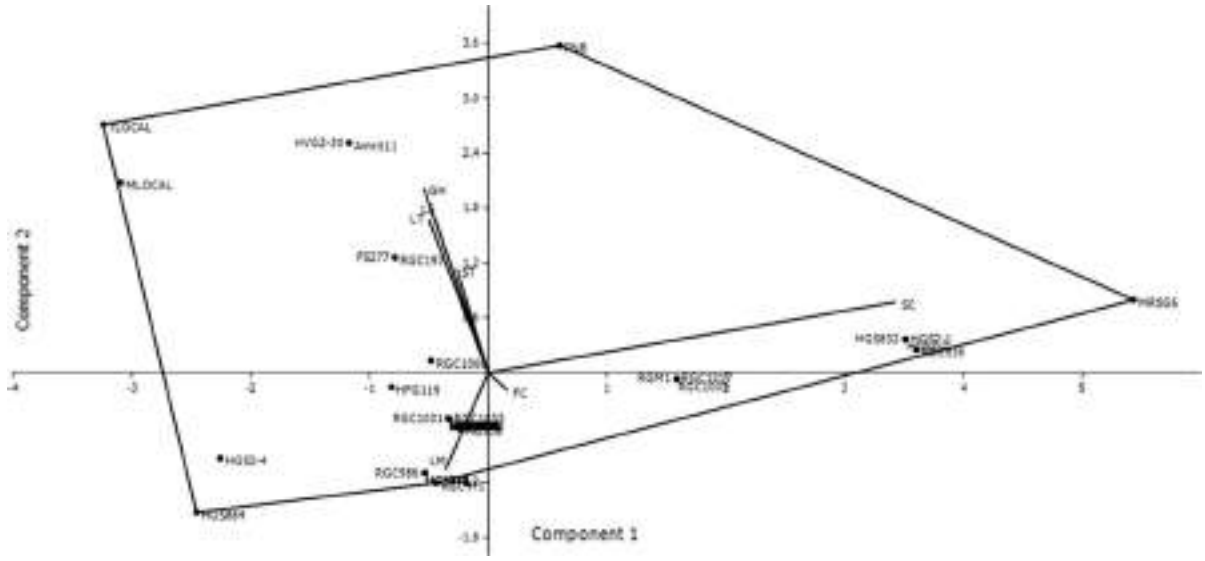


Fig. 2 Scatter plot based on PC1 and PC2 of qualitative characters

Table 7 Grouping genotypes based on qualitative traits

Clusters	Group	No of genotypes	Genotypes
I	1	2	PNB, T local
	2	3	Amrit11, HVG2-30, M local
II	1	3	RGC936, HGS16, RGC471
	2	3	RGC197, FS277, RGC1066
	3	1	HFG119
	4	30	HGS832, HGS3-2, HGS2-1, RGM1, RGC1003, RGC1017, RGC1033, RGC1031, GAU513, R local, RGC1038, MRG1786, CAZG10-2, RGC1055, HGS2-20, SRG1058, RGM2, HGS3-52, HGS75, HGS365, GAU512, RGC1002, MRGS6, HGS2-4, RGC986, HGS182, HGS870, HGS258, HGS536, HGS884

Table 8 Shannon–Weaver diversity indices (H′) of traits

Characters	SC	FC	LS	ST	LM	GH	LT	Average	s. d
H′	3.63	3.67	3.63	3.63	3.61	3.61	3.62	3.63	0.02

SC seed colour, *FC* flower colour, *LS* leaf surface, *ST* stem type, *LM* leaf margin, *GH* growth habit, *LT* leaf type

Discussion

Qualitative trait evaluation

Among the qualitative traits, the trait seed colour showed much variation (five groups), while all other traits showed only two groups. Most accessions possessed greyish seed colour, purple flowers, branched stems, bushy growth habit and narrow leaflets.

Leaflet type and leaflet surface together are easily distinguished the vegetable genotypes (HVG2-30, T local, M local, PNB and Amrit 11) from gum genotypes. The distinct feature broad and glabrous leaflet surface are helpful in identifying vegetable type, however, the trait narrow leaflet type with pubescence is associated with gum genotypes. Kumar et al. (2013) employed the above traits for varietal characterization in clusterbean. Such a unique

morphological trait can be directly employed in seed production plots to rogue out the off types (vegetable type) from the gum genotypes at field level even at initial stages of crop growth.

Five genotypes namely RGC1066, RGC197, FS277, T local and PNB were identified as non-branching types (single stem). Sivakumar (2002) and Kumar et al. (2013) also identified two genotypes namely RGC 197 and FS 277 with a single stem. Single stem genotypes are preferred as they bear more pods per cluster and also used as a preferred intercrop in cotton fields or a shade crop in ginger fields. Such a stem type also serves as a useful marker in seed production plots. Among the predominant pink colour flowers, four genotypes namely HGS 16, RGC936, RGC471 and HVG2-30 had white flower colour. Kumar et al. (2013) distinguished the genotypes based on flower colour. Flower colour aids in maintenance of purity at flowering stage which reduces contamination in seed production. The genotype MRSG6 showed brown seed with distinct grade of N 200 B which was unique and differed from the predominant grey group seeds, such seed traits useful in removing the admixture during processing stage.

High morphological variability was observed in clusterbean cultivars and respect to pubescence of the plant, pattern of branching, bearing habit, shape, size and texture of the pods, seed size and colour, and also quantity of gum in the seeds and also these traits were widely employed for characterization of clusterbean germplasm (Sultan et al. 2012; Morris 2010; Fletcher and Murphy 1998).

Correlation among qualitative traits

Pearson (1901) correlation coefficient, among the seven qualitative traits revealed that five inter correlation showed significant interaction. The trait leaflet type showed positive correlation with leaflet surface, since all vegetable genotypes namely HVG2-30, T local, M local, PNB and Amrit 11 could be easily differentiated from gum genotypes, they possesses glabrous leaflet surface and broad leaflets. Stem type and growth habit had correlation with each other as single stem type usually showed the erect growth habit, while branched stem type correlated with a bushy canopy. Hymowitz (1965) used branching pattern in clusterbean for defining the growth habit into erect and bushy classes. Paroda and Saini (1978) utilized such a kind of grouping for branch type in correlations and plant growth habit to select ideal plant type for high yield in cluster bean. The erect plant canopy with single stem occupies less crop canopy and they showed high number of pods per plant because they had more clusters per plant and more pods per cluster. High yield with limited canopy is preferred as intercrop in black cotton soils of Tamil Nadu, India.

Principal component analysis (PCA) and clustering

PCA is a powerful and the most preferred multivariate statistical technique that carries out a data reduction routine among several independent/interdependent variables while retaining those characteristics of the dataset that contribute most to its variance, by keeping lower-order principal components and ignoring higher-order ones. PCA simplifies a data set by reducing its multidimensionality to lower dimensions. PCA is done through a linear transformation that transforms the data to a new coordinate system such that the greatest variance by any projection of the data comes to lie on the first coordinate (called the first principal component), the second greatest variance on the second coordinate, and so on. These low-order principal components are further used for various comparisons among the cases.

Qualitative morphological traits are widely preferred for characterization of the germplasm because they are relatively less influenced by the environment unlike that of quantitative traits. They form discrete phenotypic classes and therefore are highly useful tools in classifying germplasm and can be predominantly assessed visually and even by novice evaluators. This is the reason why a large number of DUS (distinctiveness, uniformity and stability) characters for plant cultivar registration are defined using qualitative traits. However their numbers are often limited and the spectrum variability is relatively small (Kruskal 1978).

As a measure of reducing the multiple variables to a handful of principal components that explain hierarchically lowering levels of variance components, the PCA of the qualitative data derived from 42 genotypes, revealed at least four independent variable groupings accounting to 87.79 % of total variation (Eigen value more than 1). This is evident from the extraction of four lower-order principal components

from the PCA, which is explained by the presence of high number of independent variables. The PCA is highly effective as a data reduction tool, when variables are inter-correlated (Pearson 1901). These independent components therefore would be used for effective classification of the genotypes.

Morris (2010) applied PCA in clusterbean for identifying major source of variability among agronomical, reproductive and phenological traits in clusterbean and identified major source of variation as plant architecture i.e. stem type and growth habit contributed more amount of variation. Since the biological explanation of principal components is tricky, the best way to make sense of the PC is to find the degree of influence of each variable weight on each of the components. Primary variability was governed by the traits growth habit, leaflet surface, stem type, leaflet type, flower colour and seed colour as explained by the first four PCs.

The genotypes namely PNB, T local, HGS 884, RGC 471, MRSG6 were much diverse as observed in the biplot (Fig. 2). These diverse genotypes can be employed as distinct parents for future breeding programmes especially hybridization. There was clear grouping of vegetable genotypes along with fodder type in a single quarter and gum yielding genotypes in separate quarters. These two groups were genetically found diverse and huge variability existed between these groups. These diverse genotypes are useful to broaden the gene pool of clusterbean.

Further grouping based on standardized Euclidean distance and un-weighted paired group method using arithmetic average (UPGMA) clustering method showed two distinct clusters (Fig. 3). In this clustering, all vegetable genotypes namely PNB, T local, M local, Amrit 11 and HVG2-30 grouped into cluster I. Single stem (unbranched type) genotypes were combined into one group (group 2 of cluster II), similarly

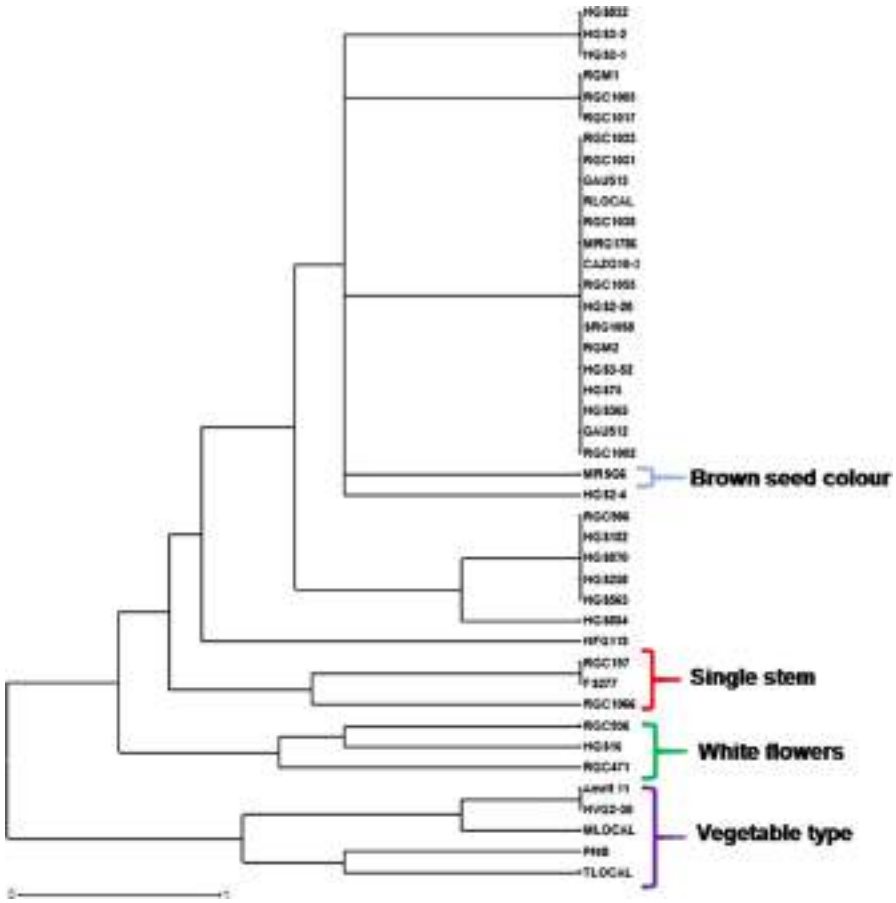


Fig. 3 Dendrogram based on UPGMA of qualitative characters

white flower colour of gum genotypes namely RGC936, HGS16 and RCG471 were grouped together (group 1 in cluster II). These distinct morphological traits can be useful for maintaining the varietal purity and also to remove the off-types and further selection programme especially pedigree breeding.

Diversity analysis

Thehighest diversity was observed in the trait flower colour based on the Shannon–Weaver diversity indices (H′) followed by seed colour, leaflet surface and stem type, leaflet type, growth habit and leaflet margin. The phenotypic dissimilarity index revealed that more diverse pairs in the order of T local versus RGC936; T local versus RGC471; T local versus MRSG6; T local versus HGS16; PNB versus RGC 471; PNB versus RGC936; PNB versus HGS16; M local versus RGC 471 and M local versus MRSG6. It is clearly evidenced that genotypes namely T local, PNB and M local were quite different from RCG936, RGC471, HGS16 and MRSG6. The same set of genotypes occupied the convex of the hull in PCA Biplot. Improvement in seed yield along with gum content can be attained by exploiting these two diverse groups.

Conclusion

This study is conducted for comparing the morphological characteristics of clusterbean accessions. The data obtained showed that germplasm resources present a wide range of diversity for morphological traits. The investigations were also very useful in choosing the most precious accessions for further breeding programmes. PCA analysis showed that genotypes namely PNB, T local, HGS 884, RGC 471, MRSG6 were much diverse, so it can be employed as a distinct parent for future breeding programme. Among the traits leaflet shape and leaflet texture are highly correlated as broad leaflets with glabrous surface mostly associated with vegetable genotypes, while narrow leaves with pubescence linked with gum genotypes. These kind of linked morphological markers are useful for gum genotype screening obviating molecular techniques. Cluster analysis revealed that vegetable genotypes (T local, PNB, M local, HVG 2-30, and Amrit11) formed as a distinct cluster from all other gum genotypes. Among the gum genotypes RGC1066, RGC197 and FS277 formed a distinct sub cluster as they possess single stem type (unbranched type), while the genotypes with white flower colour namely RGC936 and RGC471 formed a different sub cluster. The genotype MRSG6 showed brown seed with distinct grade of N 200 B which was distinctly unique and differed from the predominant grey-seeded group. In the era of Intellectual Property Rights, these morphological traits can be used as DUS (distinctiveness, uniformity and stability) characters for plant variety registration and protection. The present study clearly emphasis the power of morphological markers in characterization and utilization of plant genetic resource.

Acknowledgments The work presented in this paper is a part of the PhD work carried out by A. Manivannan at the Department of Plant Breeding and Genetics, Agricultural College and Research Institute, Madurai, Tamil Nadu, India. The support provided by Indian Council of Agricultural Research (ICAR) by awarding an ICAR-Senior Research Fellowship (ICAR-SRF) and also granting study leave to him and extending facilities at ICAR-Indian Institute of Maize Research, New Delhi, is gratefully acknowledged.

References

APEDA (2014) http://www.apeda.gov.in/apedawebsite/SubHead_Products/Guargum.htm

Department of Agriculture and Cooperation, MoA, GoI (2014). http://agricoop.nic.in

Emami MK, Sharma B (1999) Linkage between three morphological markers in lentil (*Lens culinaris* Medik.). Plant Breed 118:579–581

Fletcher R, Murphy H (1998) The new rural industries: a handbook for farmers and investors. Canberra, Australia, pp 343–346

Hammer O, Harper DAT, Ryan PD (2001) PAST: paleontological statistics software package for education and data analysis. Palaeontol Electr 4:1–9

Hymowitz T (1965) Studies on guar (*Cyamopsis tetragonoloba* (L.) Taub.). Plant breeding abstracts 36:365

Kruskal JB (1978) Factor analysis and principal component analysis: bilinear methods. In: Kruskal WH, Tannur JM (eds) International encyclopedia of statistics. The Free Press, New York, pp 307–330

Kumar S, Joshi UN, Singh V, Singh JV, Saini ML (2013) Characterization of released and elite genotypes of guar (*Cyamopsis tetragonoloba* (L.) Taub.) from India proves unrelated to geographical origin. Genet Resour Crop Evol. doi:10.1007/s10722-013-9970-7

McCormick KM, Norton RM, Eagles HA (2009) Phenotypic variation within a fenugreek (*Trigonella foenum-graecum* L.) germplasm collection. II. Cultivar selection based on

traits associated with seed yield. Genet Resour Crop Evol 56:651–661

Morris JB (2007) Swordbean (*Canavalia ensiformis* (L.) DC.) genetic resources regenerated for potential medical, nutraceutical and agricultural traits. Genet Resour Crop Evol 54:585–592

Morris JB (2008) *Macrotyloma axillare* and *M. uniflorum*: descriptor analysis, anthocyanin indexes, and potential uses. Genet Resour Crop Evol 55:5–8

Morris JB (2009) Characterization of butterfly pea (*Clitoria ternatea* L.) accessions for morphology, phenology, reproduction and potential nutraceutical, pharmaceutical trait utilization. Genet Resour Crop Evol 56:421–427

Morris JB (2010) Morphological and reproductive characterization of guar (*Cyamopsis tetragonoloba*) genetic resources regenerated in Georgia, USA. Genet Resour Crop Evol 57:985–993

NIAM (2014) An Analysis of Guar Crop in India - A study report. Jaipur, CCS NIAM, p 9

Paroda RS, Saini ML (1978) Guar breeding. In: Paroda RS, Arora RN (eds) Guar: its improvement and management. Forage Res 4:9–39

Pearson K (1901) On lines and planes of closest fit to systems of points in space. Philos Mag A 6:559–572

Purseglove JW (1981) Leguminosae. Tropical crops: Dicotyledons. Longman Group Ltd., Essex, pp 250–254

Rosso BS, Pagano EM (2001) Collection and characterization of naturalized populations of white clover (*Trifolium repens* L.) in Argentina. Genet Resour Crop Evol 48:513–517

Rubio J, Cubero JI, Martin LM, Suso MJ, Flores F (2004) Biplot analysis of trait relations of white lupin in Spain. Euphytica 135:217–224

Rumbaugh MD, Graves WL, Caddel JL, Mohammad RM (1988) Variability in a collection of alfalfa germplasm from Morocco. Crop Sci 28:605–609

Shannon CE, Weaver W (1949) The mathematical theory of communication. Univ. of Illinois Press, Urbana

Sivakumar (2002) Characterization of cluster bean cultivars using laboratory techniques and field parameters. M.Sc. thesis, Haryana Agricultural University, Hisar

Sultan M, Rabbani MA, Shinwari ZK, Masood MS (2012) Phenotypic divergence in guar (*Cyamopsis tetragonoloba* L.) landrace genotypes of Pakistan. Pak J Bot 44:203–210

Tanksley SD (1983) Molecular markers in plant breeding. Plant Mol Biol Rep 1:3–8

Ward JH (1963) Hierarchical grouping to optimize an objective function. J Am Stat Assoc 58:236–244

Zada M, Zakir N, Rabbani MA, Shinwari ZA (2013) Assessment of genetic variation in Ethiopian mustard (*Brassica carinata* A. Braun) germplasm using multivariate techniques. Pak J Bot 45(S1):583–593