Received: 23 Nov 2019 | Revised: 02 Jun 2020 | Accepted: 05 Jun 2020

Electronic Journal of Plant Breeding

Research Article

Genetic diversity assessment of extant cotton varieties based on Principal Component Analysis (PCA) and cluster analysis of enlisted DUS traits

V. Santhy[1*], K. Rathinavel[2], M. Saravanan[3], Mithila Meshram[4] and C. Priyadharshini[5]

[1,3,4]ICAR-CICR, Nagpur, Maharashtra, India
[2,5]ICAR-CICR, Regional Station, Coimbatore, Tamil Nadu, India
***E-Mail:** santhy100@gmail.com

Abstract

Morphological characterization of 47 tetraploid cotton varieties cultivated in different zones of India was carried out over two seasons. The lay out followed randomized block Design and evaluation was done using 36 DUS descriptors in two replications. The visual characters showed uniform expression within the variety for two consecutive years indicating that they were uniform and stable in expression. Eleven out of 37 traits were monomorphic among the varieties. The remaining 26 characters were used for Principal Component Analysis to find the contribution of traits towards total variability. The PCA identified a total of 10 Components with Eigen values more than 1 contributing to a cumulative 77.74 % variability. The first component (PC1) exhibited maximum variability and highly correlated with traits such as leaf shape and petal spot which are also included in the grouping characters of DUS test guideline. The scatter diagram drawn using first two principle components with highest variability as well as the hierarchical cluster analysis performed using all the ten components distinctly classified genotypes in a consistent manner. The grouping of genotypes was attributed to relatively high contribution from few characters or variables which had high positive loadings, distributed among first two components rather than small contribution from each character.

Keywords

Cotton varieties, Variety Protection, Principal Component Analysis, Genetic diversity

INTRODUCTION

Cotton is an important cash crop playing a dominant role in both agricultural and industrial economy of India. The crop is grown under diverse agro-climatic conditions with numerous popular varieties developed in each cultivated zone. With the introduction of Bt transgenic hybrid cotton in India, the cotton cultivated area and production increased significantly, which presently is, 125.86 Lakh ha and 370 Lakh bales (*www.cotcrop.org.in*) respectively. More than 90% cotton cultivated area is under private Bt hybrids which resulted in disappearance of many popular varieties which possessed various beneficial traits. The Protection of Plant Variety and Farmers Rights Act, 2001 (PPV & FRA) of India has provided opportunity to get these old popular varieties registered so as to provide a beneficial advantage for these varieties and prevent its future misuse during the development of modern genotypes. Phenotyping based on enlisted DUS (Distinctness, Uniformity and Stability) traits is a pre-requisite for the varieties to be registered under the PPV&FRA and, 37 traits have been provided in the National Test Guidelines. Understanding those traits that contribute to maximum variation among the enlisted ones would assist both in variety characterization as well as in determining the magnitude of variability in the available population which in turn is fundamental for genetic improvement.

Cluster analysis is widely used to measure the genetic distance between any two genotypes based on character or trait value and classifies the whole population into groups. Another powerful tool used in modern data analysis is Principal component analysis (PCA) which is a multivariate statistical technique capable of identifying minimum number of components that can explain the maximum variability out of the total variability (Mohammadi, 2002). The advantage of PCA is that it identifies which variable or trait is responsible for classifying the population into cluster/group. The character that contributes maximum variation is usually considered for selection. The method has been used earlier in India as well as abroad for assessing genetic diversity in cotton (Kaleri *et al*, 2015; Manivannan *et al*, 2018), Soybean (Sreenivasa *et al*, 2019), rice (Prafull *et al*, 2015), wheat Khodadadi *et al*, (2011) and finger-millet (Dosad *et al*, 2017).

This investigation was undertaken to assess the genetic diversity among the released cultivars of cotton developed from diverse agro-climatic zones of the country using enlisted morphological traits for DUS and applying PCA and cluster based methods for grouping them.

MATERIALS AND METHODS

The investigation was carried out in two seasons during *Kharif* 2014-2016, at ICAR-Central Institute for Cotton Research, Nagpur (CICR). The source material comprised 47 varieties of extant upland cotton raised in randomized block design with two replications at a spacing of 90 x 60 cm (**Table 1**). All the recommended agronomic and plant protection practices were followed to raise the crop.

For the convenience of field data recording and further comparison of traits between varieties, all the characters were recorded in terms of pre determined notes provided for each state of trait as in the national DUS test guideline for tetraploid cotton (Rathinavel *et al.* 2005). All the plant morphological characteristics were recorded under field condition from seedling till harvest at different stages of crop growth.

The presence (9) or absence (1) of pigmentation on the hypocotyls was recorded at seedling stage. The leaf characteristics such as leaf color, leaf hairiness, leaf appearance, gossypol glands, nectaries, petiole and leaf shape were recorded based on visual assessment in the fourth leaf from the top of the plant. The presence or absence of pigmentation on the stem, the hairiness on stem, flower characteristics such as bract type, petal color, petal spot, position of stigma, anther filament colouration and pollen colour were recorded at peak flowering stage The plant height and growth habit was noted in centimetre and measured at final picking stage. At the stage of boll maturity, their position as solitary/cluster, boll color, boll shape, prominence of boll tip and boll surface were observed. Boll opening pattern and fibre color were recorded at boll bursting stage. All the above traits were observed in ten plants selected randomly from each of the two replications. For boll weight, 3 bolls were harvested randomly from each of the ten plants and 10 bolls drawn from the composite bulk for fuzz color, density and seed index. Seed index was arrived by taking the average of weight 100 seeds measured in 8 replications. The ginning per cent was computed by dividing the lint weight by total kapas weight and the fibre parameters included fibre length (FL), fibre strength (FS), fibre fineness (FF), fibre uniformity (FU) and fibre maturity (FM) measured in a fully automated cotton testing instrument under ICC mode.

Principal Component Analysis (PCA) was performed on all variables simultaneously to find out the relative importance of different traits having a maximum genetic variation using XLSTAT 2019. Scatter diagram were drawn for visually assessing the components or factors of variability in the data by PAST 3 (Hammer *et al.* 2001). The factors corresponding 1 to 10 PCs were subjected to cluster analysis based on Euclidean distances using software XLSTAT 2019. A hierarchical cluster analysis for pooled data was performed using scores of dissimilarity matrix (Ward, 1963).

RESULTS AND DISCUSSION

Morphological features are traditionally used for the assessment of genetic diversity and determining genetic distances. UPOV recommends use of morphological characters for testing varietal distinctness/protection and suggests other tests, only as additional or complementary (Jones *et al*, 2013). In cotton, flower petal color, petal spot, seed fuzz color and leaf shapes have been used to distinguish popular hybrids (Hosseini, 2014). The present study on 47 varieties using 36 DUS traits revealed 10 to be monomorphic and 26 to be polymorphic with varying frequencies (**Table 2**). The above traits observed visually did not show any variation within the variety over two years of study indicating their stability and Uniformity.

The expression of traits under which majority of the genotypes were classified and showed high frequencies were palmate normal leaf (95.75%), presence of stem pigmentation (95.75%), absence of flower petal spot (95.75%), smooth boll surface (95.75%), pointed boll tip (97.87%) and medium seed fuzz density (91.49%). Three distinctive groups were observed for leaf and stem hairiness with 42 under medium (89.36%) and 2 each under dense category (4.26%). The flower pollen color revealed three classes with 31 falling under cream pollen (65.96%), 14 under yellow (29.79%) and 2 under deep yellow (4.26%). Flower petal color showed two expressions with 44 under cream (93.62%) and 3 under yellow (6.38%). Majority of the genotypes fell under semi-spreading growth habit (87.23%), whereas 4 were under spreading (8.51%) and 2 only under compact class (4.26%). Boll shape showed 23 genotypes under round (48.94%), 22 under oval (46.81%) and only two under elliptic (4.26%). Boll weight revealed variations at three levels with small (72.34%), very small (25.53%). Among the seed traits, seed fuzz density and seed index were highly varying with genotypes showing three different groups of expression in each. In a similar study on 101

Table 1. Details of extant cotton varieties

S. No.	Name	Source	Released Zone
1	CNHO12	CICR, Nagpur	Uttar Pradesh, Gujarat, Maharashtra and Madhya Pradesh.
2	Anjali	CICR, RS, Coimbatore	Rice fallow areas of Tamil Nadu, Maharastra, Rajasthan, Gujarat
3	Pratima	CICR, Nagpur	South zone
4	Suraj	CICR, RS, Coimbatore	Tamil Nadu, Karnataka, Andhra Pradesh
5	Surabhi	CICR, RS, Coimbatore	Tamil Nadu, Karnataka, Andhra Pradesh
6	Supriya	CICR, RS, Coimbatore	Tamil Nadu, Andhra Pradesh
7	Sumangala	CICR, RS, Coimbatore	High yielding hirsutum variety for for rainfed region of south zone
8	Sahana	UAS, Dharwad	Both irrigated and rain fed areas of Karnataka
9	Abhadita	UAS, Dharwad	Karnataka
10	Arogya	CICR, Nagpur	Maharastra, Gujarat, Madhya Pradesh
11	Narasimha	RARS, Nandyal	Andhra Pradesh
12	MCU5	CICR, RS, Coimbatore	Tamil Nadu, Andhra Pradesh, Karnataka, Maharashtra, Orissa
13	MCU5VT	CICR, RS, Coimbatore	Winter tract of Tamil Nadu, Andhra Pradesh
14	MCU10	TNAU, Coimbatore	Black soil of Tamil Nadu
15	MCU12	TNAU, Coimbatore	Suitable for 50’s count and moderate sucking pest tolerence
16	DHY286	Dr. PDKV, Akola	Vidharbha of Maharastra and North Andhra Pradesh
17	Bikeneri Narma	RAU, Sriganganagar	North of punjab, Haryana and Sriganganagar of Rajasthan
18	AKH8828	Dr. PDKV, Akola	Central Zone, Maharashtra
19	AKH081	Dr. PDKV, Akola	Central Zone, Maharashtra
20	Khandwa2	JNKVV, Khandwa	Nimar tract of Madhya Pradesh
21	NH545	MAU, Nanded	Drought tolerent variety for Marathwada
22	NH615	MAU, Nanded	Maharashtra
23	JK4	JNKVV, Khandwa	Maharashtra and Madhya Pradesh
24	Gcot10	NAU, Surat	Middle and south Gujarat, Maharastra
25	Gcot16	NAU, Surat	Middle and South Gujarat
26	Gcot20	NAU, Surat	Irrigated tracts of Gujarat
27	Gcot18	NAU, Surat	Suitable both for irrigated and rain-fed areas of Gujarat
28	Gcot12	NAU, Surat	Dholke area of Gujarat, Rajasthan
29	Ganganagar Ageti	RAU, Sriganganagar	Sriganganagar of Rajasthan
30	JLH168	-	-
31	F1054	PAU, Faridkot	Punjab
32	F1378	PAU, Faridkot	Punjab, Haryana, Rajasthan
33	F846	PAU, Faridkot	Punjab
34	F1861	PAU, Faridkot	Punjab
35	RS875	RAU, Sriganganagar	Rajasthan
36	RS2013	RAU, Sriganganagar	Irrigated tracts of north zone
37	RS810	RAU, Sriganganagar	Rajasthan
38	RST9	RAU, Sriganganagar	Rajasthan
39	HS6	CCSHAU, Hisar	Haryana
40	H1226	CCSHAU, Hisar	Irrigated conditions of Haryana
41	KC3	TNAU, Kovilpatti	Well suited for low rainfall areas of south zone
42	Deviraj	NAU, Surat	Central Gujarat, Decan area of Maharastra, Karnataka
43	Kanchana	RARS, Lam	Whitefly problem areas of Southern Zone
44	PKV Rajat	PDKV, Akola	Maharashtra
45	RMPBS155	UAS, Dharwad	Robust genotype responsive to inputs
46	Suvin	CICR, RS, Coimbatore	Tamil Nadu, Andhra Pradesh
47	Sujata	CICR, RS, Coimbatore	Tamil Nadu

Table 2. DUS character variables and their frequency

Variables	Scores	**Categories**	**No. of genotypes**	**Frequency %**
Hypocotyl : Pigmentation	9	Present	47	100.00
Leaf: Colour	1	Light green	5	10.64
	2	Green	42	89.36
Leaf: Hairiness	1	Sparse	3	6.38
	5	Medium	42	89.36
	9	Dense	2	4.26
Leaf: Appearance	1	Cup	16	34.04
	2	Flat	31	65.96
Leaf: Gossypol glands	9	Present	47	100.00
Leaf: Nectaries	9	Present	47	100.00
Leaf: Petiole pigmentation	9	Present	47	100.00
Leaf: Shape	1	Palmate (Normal)	45	95.75
	2	Semi-digitate	2	4.26
Plant: Stem hairiness	3	Sparse	3	6.38
	5	Medium	42	89.36
	7	Dense	2	4.26
Plant: Stem pigmentation	1	Absent	2	4.26
	9	Present	45	95.75
Plant: Height (cm)	3	Semi dwarf	15	31.92
	5	Medium tall	32	68.09
Plant: Growth habit	3	Compact	2	4.26
	5	Semi-spreading	41	87.23
	7	Spreading	4	8.51
Bract: Type	3	Normal	47	100.00
Flower: Time of flowering	5	Medium	24	51.06
(50% of plants with at least one open flower)	7	Late	23	48.94
Flower: Petal colour	1	Cream	44	93.62
	2	Yellow	3	6.38
Flower: Petal spot	1	Absent	45	95.75
	9	Present	2	4.26
Flower: Stigma	3	Embedded	29	61.70
	5	Exerted	18	38.30
Flower: Anther filament colouration	1	Absent	47	100.00
Flower: Pollen colour	2	Cream	31	65.96
	3	Yellow	14	29.79
	4	Deep yellow	2	4.26
Flower: Male sterility (Only for A and B lines)	1	Absent	47	100.00
Boll: Bearing habit	1	Solitary	47	100.00
Boll: Colour	3	Green	47	100.00
Boll: Shape (longitudinal section)	3	Round	23	48.94
	5	Ovate	22	46.81
	7	Elliptic	2	4.26
Boll: Surface	1	Smooth	45	95.75
	9	Pitted	2	4.26
Boll: Prominence of tip	1	Blunt	1	2.13
	9	Pointed	46	97.87
Boll: Opening	3	Semi-open	23	48.94
	5	Open	24	51.06
Boll: Weight of seed cotton/boll (g)	1	Very small	12	25.53
	3	Small	34	72.34
	5	Medium	1	2.13
Seed: Fuzz	3	Sparse	2	4.26
	5	Medium	43	91.49
	7	Dense	2	4.26
Seed: Fuzz colour	1	White	14	29.79
	2	Grey	33	70.21
Seed: Index (100 seed wt in gram)	3	Small	7	14.89
	5	Medium	26	55.32
	7	Bold	12	25.53
	9	Very bold	2	4.26
Ginning %	1	Very low	6	12.77
	3	Low	12	25.53
	5	Medium	9	19.15
	7	High	9	19.15
	9	Very high	11	23.40
Fibre: Colour	1	White	39	82.98
	2	Cream	8	17.02
Fibre: Length (2.5 % spanlength)(mm)	3	Medium	20	42.55
	5	Medium long	12	25.53
	7	Long	14	29.79
	9	Extra long	1	2.13
Fibre: Strength (g/tex)	3	Weak	2	4.26
	5	Medium	1	2.13
	7	Strong	37	78.72
	9	Very Strong	7	14.89
Fibre: Fineness (Micronaire value)	3	Coarse	3	6.39
	5	Medium	10	21.28
	7	Fine	29	61.70
	9	Very fine	5	10.64
Fibre: Uniformity (%)	9	Excellent	47	100.00

genotypes (different from the present study material) of cotton, Rathinavel (2018) identified petal color and seed fuzz density as the major traits expressing maximum variation among the genotypes compared to other traits. Because of the continuous efforts to breed cotton varieties for increased fibre yield and quality, it is likely that many related agronomic/taxonomic traits were not observed and might have shifted or remained unchanged. The present study of morphological diversity involving a large number of popular varieties is therefore of much significance which show us the range of variability existing for all the traits other than only the yield contributing ones.

Principal component Analysis was employed in 47 extant varieties of upland cotton to study each component or trait. PCA enables identification of plant traits that contributed most variation within a group of genotypes and has a practical application in the selection of parental lines for breeding purpose (Ahmadizadeh and Felenji, 2011). From the factor pattern and summary of the principal component analysis (Table 3) it is evident that 77.74% of the total variability among 47 genotypes of tetraploid cotton is explained by ten principal components (PCs) extracted with Eigen value >1. The principal components with Eigen values more than one is considered more informative than any single variable alone (Venujayakanth *et al*, 2017). It is therefore indicated that the identified characters within each of the ten component are highly useful and exhibited immense influence on the phenotype of the cultivars either positively or negatively. The data set revealed maximum variation for Principal Component 1 (PC1) (23.09%). The correlations between the variables and factors are called factor loadings. Three characters with maximum positive loading in PC1 included leaf shape (0.963), petal spot (0.963) and boll surface (0.963). These are also incidentally the grouping characters used in DUS testing. The other traits included in the first component were pollen color (0.678), stem hairiness (0.628), boll shape (0.602) and fibre length (0.468). This was followed by second component PC 2 contributing 9.66% variation including traits such as leaf hairiness (0.604) growth habit (0.579), seed fuzz density (0.394) and flower petal color (0.273). The third component, PC3 accounted for 7.87% variability and was influenced by characters such as plant height (0.58), ginning % (0.56) seed index (0.54) boll opening (0.30) and flower petal color (0.273).

Table 3. Principal component analysis of DUS traits in extant cotton genotypes

Characters/ PCs	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
LC	0.094	-0.419	0.248	0.249	-0.305	-0.003	-0.206	0.583	0.020	-0.213
LH	-0.648	0.604	-0.078	-0.007	0.165	0.004	0.077	-0.109	-0.128	-0.086
LA	-0.347	-0.416	0.042	0.020	-0.256	-0.325	-0.166	-0.469	0.044	0.309
LS	0.963	-0.062	0.122	-0.020	0.013	0.023	0.003	-0.113	0.006	-0.051
PSH	0.628	0.129	-0.273	-0.458	0.220	0.029	0.142	-0.013	0.118	0.142
PSP	-0.508	-0.024	-0.177	0.350	0.309	0.220	-0.083	-0.033	0.354	-0.011
PH	-0.276	-0.031	0.584	-0.088	-0.218	0.070	-0.312	0.092	-0.254	0.269
GH	-0.043	0.579	0.481	-0.339	0.001	0.029	-0.047	-0.001	-0.216	-0.292
FLO	0.316	0.399	-0.065	0.266	0.323	-0.031	-0.289	0.217	-0.124	0.338
FPC	-0.115	-0.087	0.273	0.115	-0.062	0.608	0.074	-0.023	-0.083	0.417
FPS	0.963	-0.062	0.122	-0.020	0.013	0.023	0.003	-0.113	0.006	-0.051
FS	0.248	-0.553	0.014	-0.152	0.342	0.247	0.049	-0.070	-0.149	-0.312
FPOC	0.678	0.319	-0.162	0.062	-0.169	0.006	-0.196	-0.154	0.073	-0.088
BSh	0.602	0.289	-0.079	0.359	0.122	-0.117	-0.032	0.082	0.277	0.039
BS	0.963	-0.062	0.122	-0.020	0.013	0.023	0.003	-0.113	0.006	-0.051
BT	0.043	-0.057	0.190	0.120	0.543	-0.295	0.502	0.146	-0.136	0.205
BO	-0.340	-0.096	0.307	-0.081	-0.173	-0.500	0.259	-0.318	0.119	-0.178
BW	-0.152	0.080	-0.218	-0.092	-0.279	0.622	0.068	-0.233	0.273	-0.101
SF	-0.715	0.394	-0.186	-0.158	0.066	-0.010	0.136	0.188	0.145	-0.046
SFC	-0.168	-0.345	0.213	0.656	0.183	0.142	0.329	-0.200	0.027	-0.075
SI	-0.040	0.300	0.549	-0.045	-0.089	-0.007	-0.082	-0.094	0.477	0.124
G%	0.066	0.038	0.562	-0.337	0.352	0.278	0.186	-0.113	0.054	0.091
FC	-0.017	-0.480	-0.241	-0.596	-0.008	-0.106	0.088	0.255	0.285	0.318
FL	0.468	0.371	-0.003	0.263	-0.299	-0.144	0.209	-0.109	0.023	0.240
FS	0.169	0.120	0.266	0.049	-0.403	0.036	0.536	0.403	0.248	-0.069
FF	0.109	0.093	-0.366	0.019	-0.462	0.155	0.457	-0.038	-0.433	0.137
Eigen values	6.00	2.51	2.04	1.79	1.68	1.48	1.37	1.22	1.09	1.03
Variability (%)	23.09	9.66	7.85	6.87	6.48	5.71	5.28	4.68	4.18	3.95
Cumulative (%)	23.09	32.75	40.60	47.47	53.95	59.66	64.94	69.62	73.80	77.74

LC-Leaf colour, LH-Leaf hairiness, LA-Leaf appearance, LS-Leaf shape, PSH-plant stem hairiness, PSP-plant stem pigmentation, PH-plant height, GH-growth habit, FLO-50% flowering, FPC-flower petal colour, FPS-flower petal spot, FS-flower stigma, FPOC-flower pollen colour, BSh-boll shape, BS-boll surface, BT- boll tip, BO-boll opening, BW-boll weight, SF-seed fuzz, SFC-seed fuzz colour, SI-seed index, G%-ginning per cent, FC-fiber colour, FL-fiber length, FS-fiber strength, FF-fiber fineness

The first three components alone explained 40.60% of total variability and remaining components together showed only 37.14% variation. It is a common practice to retain all those components that together are contributing towards 70% to total variability (Jolliffe and Cadima, 2016). However, the emphasis in PCA is generally on the first few PCs, though there are circumstances in which the last few may also be of interest as observed in the present study (Jolliffe, 2002). The fourth component (PC4) included traits such as seed fuzz color (0.65) and stem pigmentation (0.35) while the variation in component five (PC 5) was contributed mainly by boll prominence of tip (0.54) and flower stigma exsertion (0.34). The characters in sixth component were boll weight (0.62) and petal color (0.60) and seventh component was mainly influenced by fibre strength (0.53) and fibre fineness (0.45). The leaf color (0.58) and fibre colour (0.25) had the maximum loadings in eighth component. Eight useful Principal Components including those variables related to yield performance and fiber quality has been identified earlier in cotton by Rathinavel, (2018). PCA has been extensively used for investigating variation level among different traits, during the selection of parents for developing varieties tolerant to ClCuD in cotton (Javed *et al*, 2017).

Scatter plot also named polygon as generated through the principal components analysis illustrated the genetic pattern of extant upland cotton genotypes (Fig. 1). Among the ten components, two components (PCA1 and 2) that exhibited the highest variability were taken for scatter plot analysis. The genotypes which clogged at the vertex of the polygon are farthest from point of origin, hence more diverse. The genotypes *viz.,* Suvin, Sujatha, Kanchana, Arogya, G.Cot.10, Suraj, G.Cot.12 and DHY 286 were found at different vertices of the polygon and distinct from other genotypes. The two genotypes, Suvin and Sujatha possessed traits specific to *G. barbadense* and hence were distinctly placed in a separate cluster and farthest from the other genotypes of *G. hirsutum*. These two were clearly distinguished by presence of semi digitate leaf, presence of petal spot, deep yellow pollen color, elliptic boll shape with pitted surface, sparsely fuzzy seed and extra long fibers. Among *G. hirsutum*, G.Cot.12 and DHY 286 together were placed distinct and possessed densely haired leaves and stem. Two other varieties of *G. hirsutum,* Suraj and Kanchana were individually placed as distinct at separate vertices and are known to possess longer fibres and sparse pubescence of leaf/stem respectively. These genotypes were placed distinctly based on variables included in the first two Principal Components and are very much useful for the breeding programme/ variety characterization. The utilization of scatter plot to find out most phenomenal and divergent genotypes was done earlier by Rehman *et al.,* (2015) and Manivannan *et al* (2018) in cotton.

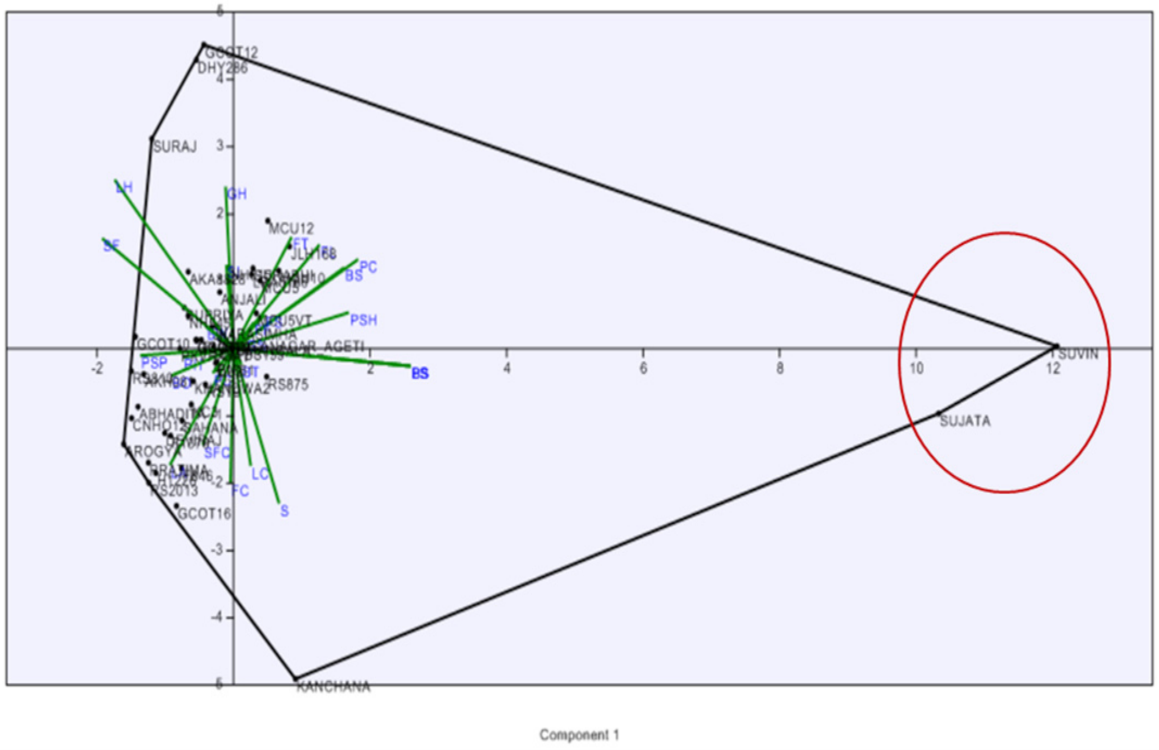

Fig. 1. Scatter diagram of Biplot showing cotton varieties

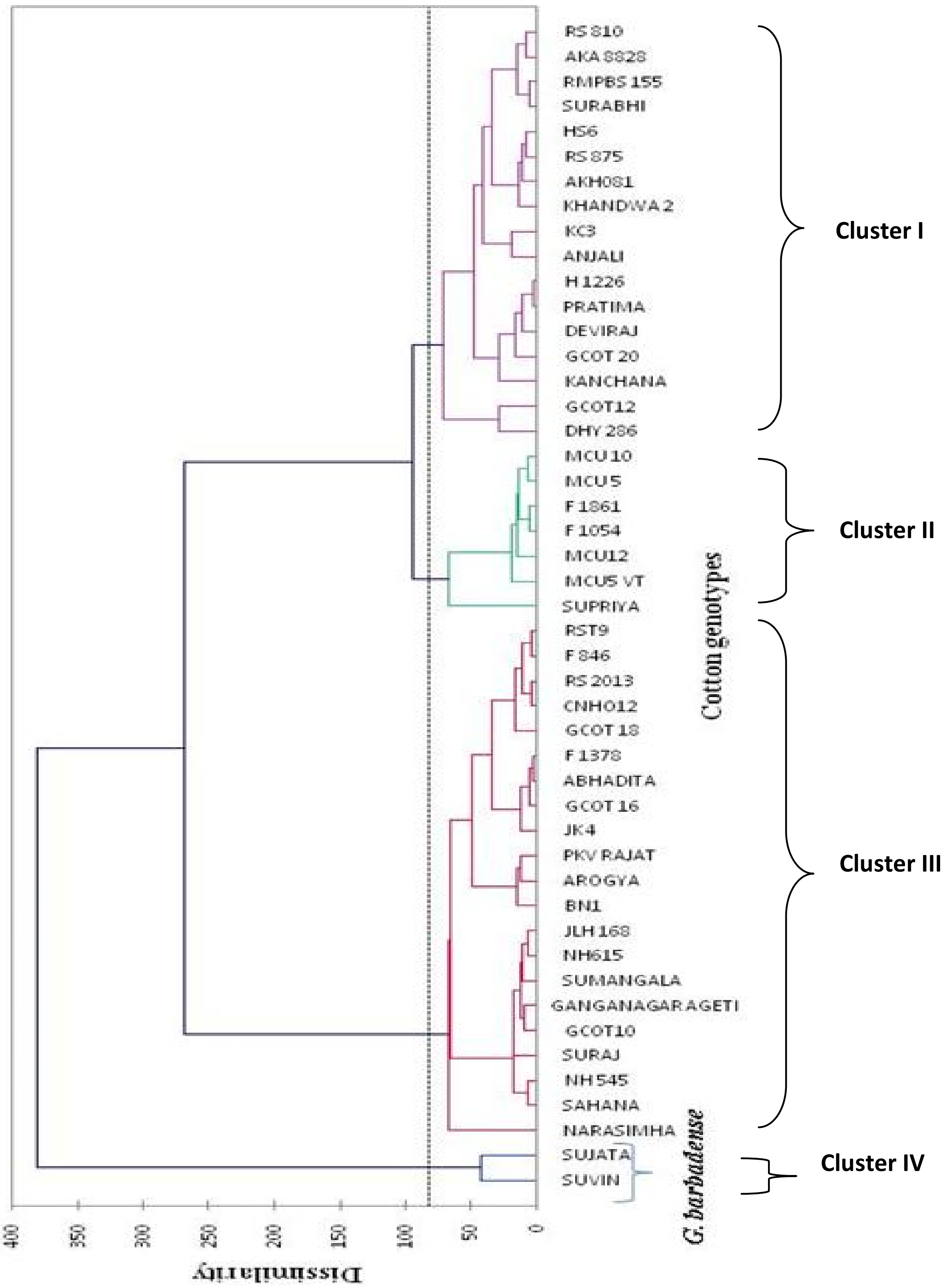

Fig. 2 Dendrogram showing the clusters of extant cotton varieties

The factors corresponding to 10 Principal Components were subjected to cluster analysis using scores of dissimilarity matrix. The hierarchical cluster analysis grouped the 47 genotypes of upland cotton into three major clusters based on Ward's minimum distance variance method (**Fig.2; Table 4**). Genotypes of *G. barbadense* formed a distinct group which are based on characters that contributed maximum to the Principal Component 1. They were placed farthest from remaining cultivars. According to Chahal and Gosal (2002), characters with largest absolute value closer to unity within the first principal component influence the clustering more than those with lower absolute value closer to zero. The clusters were numbered according to the number of varieties they encompassed. Accordingly, Cluster I comprised of 21 genotypes with 5 sub-clusters. The genotypes, Suraj and Narasimha were placed distinct within this cluster. The cluster II included 17 genotypes with 5 sub-clusters. The varieties, DHY 286 and G Cot 12 were closer (both densely haired) and formed one sub-cluster within this. Seven genotypes were grouped in cluster III which was distributed into two sub-clusters. The variety, Supriya, known for its blunt boll tip was placed distinctly from the rest. The results of cluster analysis were closely in line with those of the spread out plot results. The genotypes showing distinctness in scatter diagram were retained as distinct in cluster analysis also. The major differentiation of genotypes into different clusters is attributed to relatively high contribution from few characters or variables distributed among first three components rather than small contribution from each character.

Table 4. Inter cluster distance of extant cotton varieties

Clusters	I	II	III	IV
I	-	4.86	6.21	14.34
I		-	4.37	14.59
III			-	14.01

The morphological characterization was performed in 47 popular cotton varieties using 36 traits given in the DUS guidelines of tetraploid cotton. In the Principal Component Analysis performed using 26 varying traits, the traits used for grouping of cotton varieties such as leaf shape, pollen color, boll shape, fiber length, petal spot, stem hairiness have contributed more towards first component and this variance will be highly influential for the determination of distinctiveness. In the present day regime of Intellectual Property Rights, the most varying morphological traits are highly useful for DUS characterization for plant variety registration and protection of breeders' Rights. The *G. barbadense* genotypes Suvin and Sujatha remained distinct from rest of the genotypes. The information on genotypes and distinct traits contributing for more variation shall enable breeding of newer high yielding cultivars with distinct identified traits so that varieties are better protected in future.

REFERENCES

Ahmadizadeh, M and Felenji, H. 2011. Evaluating diversity among potato cultivars using agro-morphological and yield components in fall cultivation of Jiroft Area. *Am. Eurasian J. Agric. & Environ. Sci.* **11** (5): 655-662.

Chahal, G. and Gosal, S.S. 2002. Principles and procedures of Plant Breeding: Biotechnological and conventional approaches. Narosa Publishing House, New Delhi, **Vol. 21**, 64-89.

Dosad, S., Singh, M., Pandey P.K and Chawla, H.S. 2017. Assessment of distinctiveness, uniformity and stability of finger Millet (*Eleusine coracana*) varieties. *Indian J. Plant Genet. Resour.,* **30** (1): 55–60. [Cross Ref]

Hammer, O., Harper, D.A.T and Ryan, P. D. 2001. PAST: Paleontological statistics software package for education and data analysis. *Palaeontol. Electr.*, **4**: 1–9.

Hosseini, G. 2014. Estimation of genetic parameters for quantitative and qualitative traits in cotton cultivars (*Gossypium hirsutum* L. & *Gossypium barbadense* L.) and new scaling test of additive– dominance model. *Journal of Plant Mol. Breed.*, **2**(1)**:** 45-63.

Javed, M., Hussain, S.B and Baber, M. 2017. Assessment of genetic diversity of cotton genotypes for various economic traits against cotton leaf curl disease (CLCuD). *Genet. Mol. Res.,* **16** (1). [Cross Ref]

Jolliffe, I.T and Cadima J. 2016. Principal component analysis: a review and recent developments. *Phil. Trans. R. Soc. A.,* **374**: 20150202. [Cross Ref]

Jolliffe, I. 2002. Principal Component Analysis. 2nd Edition, Springer, New York.

Jones, H., Norris, C., Cockram, J and Lee, D. 2013. Variety protection and plant breeders' rights in the DNA era In: Diagnostics in Plant Breeding eds: Libberstedt T. and R K Varshney *Springer Science.* [Cross Ref]

Kaleri, A.A., Rajput, S.Y., Kaleri, G.A., Kaleri, M.K and Marri, J. 2015. Analysis of Genetic diversity in genetically modified and non-modified cotton (*Gossypium hirsutum* L.,) genotypes. *IOSR Journal of Agriculture and Veterinary Science,* **8** (12): 70-76.

Khodadadi, M., Fotokian, M.H and Miransari, M. 2011. Genetic diversity of wheat (*Triticum aestivum* L.) genotypes based on cluster and principal component analyses for breeding strategies. *Australian Journal of Crop Sciences,* **5**(1):17-24.

Manivannan, A., Mohan, P., Saravanan, M and Waghmare, V.N. 2018. Morphological characterization of asiatic cotton (*G. arboreum*) germplasm of India. *Electronic Journal of Plant Breeding,* **9** (2): 673-681. [Cross Ref]

Mohammadi, S.A. 2002. Statistical methods in genetics. Paper presented at the 6th International conference of Statistics, University of Tarbiat modares, Iran, 26-28 August 2002.

Prafull Kumar, Abhinav Sao, A. K. Gupta and Manish Kumar 2015. Principal component analysis for assessment of genetic diversity in upland paddy for Bastar plateau. *Electronic Journal of Plant Breeding*, **6** (4): 1052- 1059

Rathinavel, K. 2018. Principal component analysis with quantitative traits in extant cotton varieties (*Gossypium hirsutum* L.) and parental lines for diversity. *Current Agriculture Research Journal,* **6** (1): 54-64. [Cross Ref]

Rathinavel, K., Manickam, S and Sabesh, M. 2005. DUS testing in Cotton. Training Manual on DUS Test in Cotton with reference to PPV & FR legislation, 2001. Pg. 1-141.

Rehman, S.U., Abid, M.A., Malik, W., Hanif, C.M., Bilal, M and Qanmber, G. 2015. Exploitation of germplasm for plant yield improvement in cotton (*G. hirsutum* L.). *J. Green Physiol. Gen. Genom,* **1**(1).

Sreenivasa, V., Lal, SK., Talukdar, A., Kiran Babu, P., Mahadeva Swamy, H.K., Rathod D.R., Yadav, R.R., Poonia, S., Bhat, K.V and Viswanathan, C. 2019. Evaluation of soybean germplasm lines for agro-morphological traits and terminal drought tolerance. *International Journal of Current Microbiology and Applied Science,* **8**(4):105-127. [Cross Ref]

Venujayakanth, B., Dudhat, A.S and Swaminathan, B. 2017. Assessing crop genetic diversity using principle component analysis: A Review. *Trends in Biosciences,* **10** (2): 523-528.

Ward, J.H. 1963. Hierarchical grouping to optimize an objective function. *J. American Statistical Assoc*, **58** (301): 236-244. [Cross Ref]