PLANT SCIENCE TODAY
ISSN 2348-1900 (online)
Vol 12(1): 1-10
https://doi.org/10.14719/pst.4160

HORIZON
e-Publishing Group
HePG

RESEARCH ARTICLE

Unveiling genetic variation in Egyptian cotton (*Gossypium barbadense* L.) germplasm: A combined approach of morphological characterization and multivariate analysis

Shanmugam P Harini[1], N Premalatha*[2], A Subramanian[2], N Manikanda Boopathi[3] & K Guruswamy[4]

[1]Department of Genetics and Plant Breeding, Tamil Nadu Agricultural University, Lawley Road, Coimbatore 641 003, Tamil Nadu, India

[2]Department of Cotton, Centre for Plant Breeding and Genetics, Tamil Nadu Agricultural University, Lawley Road, Coimbatore 641 003, Tamil Nadu, India

[3]Department of Plant Biotechnology, Centre for Plant Molecular Biology and Biotechnology, Tamil Nadu Agricultural University, Lawley Road, Coimbatore 641 003, Tamil Nadu, India

[4]Department of Food Processing, Tamil Nadu Agricultural University, Lawley Road, Coimbatore 641 003, Tamil Nadu, India

*Email: npremalatha@gmail.com

ARTICLE HISTORY

Received: 23 August 2024
Accepted: 24 September 2024

Available online
Version 1.0 : 20 December 2024
Version 2.0 : 01 January 2025

Additional information

Peer review: Publisher thanks Sectional Editor and the other anonymous reviewers for their contribution to the peer review of this work.

Indexing: Plant Science Today, published by Horizon e-Publishing Group, is covered by Scopus, Web of Science, BIOSIS Previews, Clarivate Analytics, NAAS, UGC Care, etc
See https://horizonepublishing.com/journals/index.php/PST/indexing_abstracting

CITE THIS ARTICLE

Harini P S, Premalatha N, Subramanian A, Boopathi N M, Guruswamy K. Unveiling genetic variation in Egyptian cotton (*Gossypium barbadense* L.) germplasm: A combined approach of morphological characterization and multivariate analysis . Plant Science Today. 2025; 12(1): 1-10. https://doi.org/10.14719/

Abstract

Cotton (*Gossypium* sp.) is a commercial crop that is predominantly grown in tropical and subtropical regions, with India emerging as the largest producer globally. This study focused on the genetic diversity and morphological characterization of *Gossypium barbadense* L., the second most cultivated cotton species. Comprehensive data encompassing qualitative and quantitative traits, spanning morphological attributes to yield-contributing parameters were documented and analyzed using appropriate statistical analysis. The findings indicate the significance of specific attributes, *viz.,* the number of monopodia and sympodia per plant, boll weight and fiber fineness, which emerged as highly reliable markers for selection. These traits exhibit elevated genotypic and phenotypic coefficients of variation, along with substantial heritability and genetic advance as a percent of the mean. Based on *per se* performance, genotypes SBS 66, SBS 79 and EC 98254 will serve as a source for the selection of parents for improving traits *viz.,* boll weight, upper half mean length, fibre fineness and single plant yield. Principal component analysis divided the total variance into 14 components and cluster analysis segregated the germplasm into 5 distinct clusters, highlighting the substantial genetic variation and the genotypes 3530, CBS 202, ERB 4488 and EGYPTIAN, which can be used as parents in the future breeding programme.

Keywords

DUS characterization; *G. barbadense* ; multivariate analysis; variability

Introduction

Cotton (*Gossypium* sp.) is a commercial crop that is grown predominantly in the tropics and subtropics. India is the major producer of cotton, producing 5.31 million metric tonnes, or 21 % of the world's total cotton production. India is the primary habitat of all 4 major cotton species: *G. hirsutum, G. arboreum, G. herbaceum* and *G. barbadense*. *G. hirsutum* is the source of 94 % of hybrid cotton produced in India. Egyptian cotton (*G. barbadense* L.) is grown for its extra-long, sturdy and finest fiber (1). It is the second most cultivated species and accounts for less than 2 % of all cultivated cotton in the world. The southern regions of India are the primary producers of long and extra-long staple (ELS) cotton and demand for this product continues to expand annually. India currently produces 2 lakh ha of ELS cotton. However, only approximately 4 lakh bales are available to meet the country's requirements and the demand is approximately 9 lakh bales.

In the case of Egyptian cotton, the exploration of genetic variation has become even more critical, given its economic significance as an extra-long staple cotton. The genetic improvement of crop plants relies heavily on the availability of diverse germplasm with desirable traits. This diversity is essential for breeding programs aimed at enhancing yield and fiber quality, resistance to pests and diseases and adaptability to varying environmental conditions, as it provides a broader genetic base. The value of germplasm can be determined only when it is properly characterized. Characterization and evaluation of germplasm and quantification of genetically diverse genotypes are indispensable for the pragmatic use of plant genetic resources and for determining evolutionary relationships (2). A successful breeding program depends on the complete knowledge and understanding of the genetic diversity within and among genetic resources of the available germplasm, enabling plant breeders to choose parental sources that will generate diverse populations for selection.

In addition to deciphering the genetic diversity of crop germplasm utilizing quantitative traits, variations in the germplasm collections could also be investigated by employing qualitative traits. Researchers have employed morphological, chemical, and biochemical traits in the identification of genotypes in *Gossypium* sp. (3). In cotton, genotypic and phenotypic variation was studied by many researchers for traits like plant height, number of sympodia and monopodia (4). The variability and heritability are utilized to acknowledge the traits suitable for genetic improvement. Genetic variability between genotypes has been extensively explored through principal component analysis (PCA), correlation analysis and clustering approaches (5).

Hence, the present investigation was conducted with the objectives of acknowledging the traits suitable for Distinctiveness, Uniformity and Stability (DUS) testing and assessing the genetic variation and diversity existing in the *G. barbadense* L. accessions through multivariate analysis and understanding the associations among them.

Materials and Methods

Plant materials

A total of 100 accessions of *Gossypium barbadense* L. germplasm along with 4 standard check varieties, namely suvin, CO 18, CCB-51 and CCB–143B, were raised in Augmented Block Design I in uniform environmental conditions at the Department of Cotton, Centre for Plant Breeding and Genetics, Tamil Nadu Agricultural University, Coimbatore, India, during the winter of 2023 and the recommended agronomic practices were followed to enhance crop growth throughout the cropping season (6).

Data observation

Data concerning 18 qualitative attributes at the appropriate development phase, namely leaf colour, leaf hairiness, leaf shape, leaf appearance, petiole pigmentation, flower petal colour, petal spot, anther colour, filament colour, position of stigma, stem pigmentation, bearing habit, boll shape, boll pitting, boll size, boll colour, boll tip and bract type were documented employing the DUS criteria (7). Characteristics such as leaf colour, leaf hairiness, leaf petiole pigmentation, leaf shape and leaf appearance were recorded based on a visual inspection of the plant's 4th leaf from the top at the stage of 50 % flowering. The presence or absence of pigmentation on the stem was noted at the 50 % flowering stage, as were flower characteristics such as petal colour, petal spot, the position of stigma and the coloration of the anther and filaments. During the boll bursting stage, boll colour, boll shape, prominence of boll tip and boll pitting were documented.

Data on various seed cotton yield contributing traits (quantitative traits), *i.e.,* days to 50 % flowering, plant height (cm), number of monopodia/plants, number of sympodia/plant, boll weight (g), lint index (g), seed index (g), ginning out turn (%) and fiber quality parameters like Upper Half Mean Length (UHML) (mm), uniformity index (%), micronaire (µg/inch), fiber strength (g/tex) and elongation percent (%) were measured using High Volume Instrumentation (Model: USTER HVI classic 900).

Statistical analysis

The data collected is analysed using an analysis of variance approach (6). In this investigation, the phenotypic coefficient of variation (PCV) and genotypic coefficient of variation (GCV) were computed (8). Simultaneously, we determined heritability in the broad sense (h^2) and genetic advance as a % of the mean (GAM), adhering to the formulations (9). Correlation studies were performed using R Studio version 2024.04.1. Additionally, multivariate analyses such as principal component analysis (PCA) and Agglomerative hierarchical cluster analysis using the ward's linkage approach were performed utilizing the statistical package Minitab version 17.

Results

Qualitative traits are less affected by the environment than quantitative characteristics; they are significantly more useful in characterizing germplasm (10). Observations on 18 qualitative traits were recorded for 100 genotypes and the scores were analyzed. Variation was absent for the characters, *viz.,* anther colour, filament colour, the position of stigma, bearing habit, boll colour, boll tip, boll pitting and bract types. Meanwhile, the remaining 10 attributes have been identified to show variation under study. The majority of genotypes contained green leaves (74 %), and the remaining 26 genotypes were found to have light green leaves. Three categories of leaf hairiness were observed, namely: sparse hairs (93 %), medium hairs (2 %) and dense hairs in the 5 genotypes. In leaf appearance, the majority of leaves were observed to have a flat appearance (97 %) and the remaining 3 genotypes were observed to have a cup appearance. The genotypes were categorised as 2 distinct classes based on their leaf shape: 60

genotypes showed semi-digitate leaves, whereas 40 genotypes were identified as having palmate leaves. The trait of petiole pigmentation was documented in 95 genotypes (95 %) and absent in 5 genotypes, *viz.,* RRB 224, SBS (YF), SBS 49, 3530, 5391–T.

In 85 genotypes (85 %), the trait stem pigmentation was present and in 15 genotypes (15 %), it was not present. Regarding flower petal colour, yellow flowers were documented in 75 genotypes (75 %) and cream-coloured flowers were found in 25 genotypes. The trait petal spot was found in 94 genotypes (94 %) and absent in 6 genotypes: EC 136452/B, ERB 4488, GIZA, EC 101783, GIZA 14647 and Marred. A higher frequency of elliptical bolls (57 %) was observed in 57 genotypes, whereas ovate-shaped bolls were observed in 37 genotypes (37 %) and round-shaped bolls were observed in 6 genotypes. Regarding boll size, 3 categories of boll size were observed, among which 42 genotypes were observed with medium-sized bolls (42 %), 15 genotypes with large-sized bolls and 43 genotypes with small-sized bolls. (Table 1.) All the genotypes in *G. barbadense* were observed with exerted stigma, yellow anther, cream coloured filament, green bolls, pitted bolls with pointed boll tip, normal bract and solitary bearing habit (Fig. 1).

Per se performance

The selection process predominantly relies on the evaluation of the mean performance of genotypes. Based on the *per se* performance, regarding the identification of donors for a specific trait, the highest performer of particular traits is considered. (Table 2). Among the 100 accessions, a greater number of sympodia per plant was recorded by Monspi serrate (26.8), boll weight was recorded by EC 98254 (4.78 g), highest ginning outturn was observed by EC 131979 (44.64 %), high upper half mean length was recorded for SBS 66 (34 mm), higher fibre strength by SBS 66 (39.7 g/tex), uniformity index by G-122A-45 (84.5 %), elongation % was highest for Barbados (6.2 %) and highest single plant yield was observed for SBS 79 (106.8 g). (Table 3)

Analysis of variance

The outcomes of the analysis of variance implied significant differences among the 100 germplasm accessions in terms of the following traits: days to 50 % flowering, plant height, number of monopodia/plants, number of sympodia/plant, boll weight, lint index, seed index, ginning out turn, single plant yield, upper half mean length (UHML), fiber fineness and elongation %. These findings elucidated in Tables 4 and 5 highlight significant variability existing among the accessions.

Table 1. Morphological description of *Gossypium barbadense* accessions.

Sl. No.	Descriptor traits	Categories	Scores	Type of assessment	Number of genotypes	Frequency %
1	Leaf colour	Green	1	VS	74	74
		Light green	2		26	26
2	Leaf hairiness	Sparse	1	VS	93	93
		Medium	5		2	2
		Dense	9		5	5
3	Leaf appearance	Cup	1	VS	3	3
		flat	2		97	97
4	Leaf shape	Palmate	1	VS	40	40
		Semi - digitate	2		60	60
5	Petiole pigmentation	Absent	1	VS	5	5
		Present	2		95	95
6	Stem pigmentation	Absent	1	VS	15	15
		Present	9		85	85
7	Petal colour	Cream	1	VS	75	75
		Yellow	2		25	25
8	Petal spot	Absent	1	VS	6	6
		Present	2		94	94
9	Boll shape	Round	3	VG	6	6
		Ovate	5		37	37
		Elliptical	7		57	57
10	Boll size	Medium	1	VG	43	43
		Large	2		42	42
		Small	3		15	15

VG – Visual assessment by a single observation of a group of plants or parts of plants, **VS** – Visual assessment by a single observation of individual plants or parts of plants.

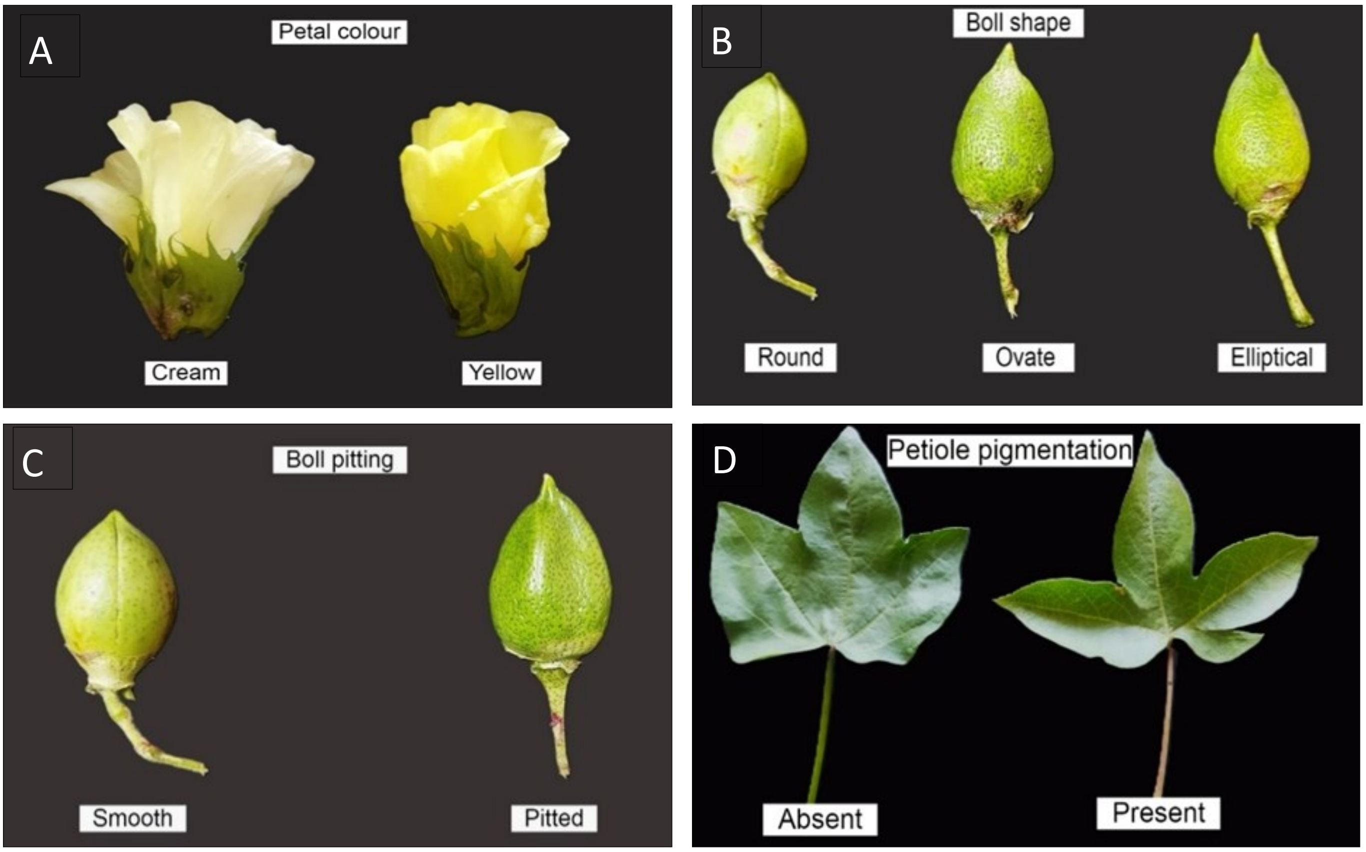

Fig. 1. DUS characterization of *G. barbadense* L. germplasm: (**A**) petal colour; (**B**) boll shape; (**C**) Boll pitting; (**D**) Petiole pigmentation.

Table 2. Potential donors for yield and fibre quality traits.

Characters	Potential accessions
Days to 50 % flowering	3530 (59 days), TCB 472/5 (60 days), SIA – 5 (60 days)
Plant height (cm)	Giza 7/A (127.06), EC 97639 (117.06), PSH (107.5)
Number of sympodia/plant	Monspi serrate (26.8), EC 27618 (26.3), Nevis (25.3)
Boll weight (g)	EC 98254 (4.78), Giza 7 (4.67), Giza 7/A (4.65)
Ginning outturn (%)	EC 131979 (44.64), Menoubi (42.68), EC 97619 (42.46)
Lint index (g)	18/4 (6.4), EC 101784 (6.2)
Seed index (g)	Egyptian (12.6), EC 101784 (11.4)
Upper half mean length (mm)	SBS 66 (34), EC 97620 (33.2), SBS (YF) (33.2)
Fibre strength (g/tex)	SBS 66 (39.7), EC 97633 (38.6), Barbados (38.5)
Uniformity index (%)	G-122A-45 (84.5), EC 4530 (84.3), EC 97634 (84)
Elongation percentage (%)	Barbados (6.2), BCS 10-91, IBSI (53)
Single plant yield (g)	SBS 79 (106.8), CBS 202 (104.1), G-122A-45 (86.2)

Variability studies

The data on genetic parameters, including phenotypic coefficient of variation (PCV), genotypic coefficient of variation (GCV), broad sense heritability (h^2) and genetic advance as a % of the mean (GAM), for 14 studied traits are furnished in Table 5. The differences between GCV and PCV for traits observed in *G. barbadense* accessions were minimal for each trait. For the following traits, high GCV with high PCV were noted: number of monopodia per plant, number of sympodia per plant, boll weight and fiber fineness. Moderate PCV and moderate GCV were documented for traits such as plant height, seed index, lint index and fiber strength. Low PCV coupled with low GCV was recorded for attributes such as days to 50 % flowering, upper half mean length, elongation % and uniformity ratio. Ginning outturn was noted with moderate PCV coupled with low GCV, whereas high PCV with moderate GCV was noted for single plant yield.

Estimates of heritability (h2) in the broad sense and genetic advance as percent of mean

Substantially high heritability was recorded for seed index, followed by days to 50 % flowering, fiber fineness, plant height, number of sympodia per plant, boll weight, number of monopodia per plant, lint index, upper half mean

Table 3. Correlation matrix among the 14 characters studied.

Traits	D5F	PH	NMP	NSP	BW	GOT	LI	SI	SPY	UHML	MIC	FS	UI	EL
D5F	1.000													
PH	0.10	1.00												
NMP	-0.04	0.17	1.00											
NSP	0.05	-0.18	-0.23*	1.00										
BW	0.06	-0.09	0.02	0.10	1.000									

GOT	0.06	0.04	-0.08	0.001	-0.11	1.00								
LI	-0.28**	-0.04	0.13	-0.001	0.04	0.12	1.00							
SI	-0.24*	-0.01	0.10	-0.12	0.052	-0.23*	0.66**	1.00						
SPY	-0.30**	-0.15	-0.008	0.05	0.08	-0.04	0.11	-0.03	1.00					
UHML	-0.03	-0.12	0.05	-0.10	-0.08	0.04	-0.10	0.06	-0.10	1.00				
MIC	-0.16	-0.03	-0.15	-0.04	-0.03	0.16	0.23*	0.07	0.02	-0.43**	1.00			
FS	0.002	-0.06	0.11	-0.11	0.02	-0.14	-0.02	0.14	-0.04	0.7**	-0.65**	1.00		
UI	0.01	0.09	0.11	-0.10	0.002	-0.22*	-0.01	0.09	0.05	0.20*	-0.17	0.33**	1.00	
EL	-0.01	-0.12	0.03	-0.12	-0.02	-0.009	0.05	0.08	0.01	0.65**	-0.39**	0.78**	0.17	1.00

*Indicates significance at 5 % level, ** Indicates significance at 1 % level, **PH** – plant height, **NMP** – number of monopodia per plant, **NSP** - number of sympodia per plant, **BW** – boll weight, **GOT** – ginning outturn, **LI** – lint index, **SI** – seed index, **SPY** – single plant yield, **UHML** – upper half mean length, **Mic** – Fibre fineness, **UI** – uniformity ratio, **EL** – elongation %.

Table 4. Analysis of variance (ANOVA) for yield and fibre quality traits.

Source of variation	*df*	D5F	PH	NMP	NSP	BW	GOT	LI
Treatments	103	42.97	300.79	0.4930	30.97	0.5269	14.693	0.70
Checks	3	7.5	243.24	0.1946	45.14	0.0112	26.9	0.15
Test entries	99	43.64**	225.01**	0.507**	29.69**	0.544**	14.12**	0.32**
Check vs test	1	83.12	7975.74	0.0036	114.96	0.3807	34.33	39.2
Error	4	0.75	9.68	0.0247	1.285	0.0288	1.43	0.01

*Indicates significance at the 5 % level, ** Indicates significance at the 1 % level, ***df***- degrees of freedom, **D5F** – days to 50 % flowering, **PH** – plant height, **NMP** – number of monopodia per plant, **NSP** - number of sympodia per plant, **BW** – boll weight, **GOT** – ginning outturn, **LI** – lint index.

Table 5. Analysis of variance (ANOVA) for yield and fibre quality traits.

Source of variation	*df*	SI	SPY	UHML	Mic	FS	UI	EL
Treatments	103	2.1826	415.26	5.57	0.62	19.21	2.49	0.06
Checks	3	0.188	1007.6	2.90	0.02	35.49	1.34	0.09
Test entries	99	1.36**	354.79	3.95**	0.65**	17.71	2.41	0.06**
Check vs test	1	89.28	4625.4	173.0	0.031	118.8	13.89	0.02
Error	4	0.007	196.47	1.28	0.026	7.556	1.646	0.008

*Indicates significance at the 5 % level, ** Indicates significance at the 1 % level, ***df***- degrees of freedom, **SI** – seed index, **SPY** – single plant yield, **UHML** – upper half mean length, **Mic** – Fibre fineness, **UI** – uniformity ratio, **EL** – elongation %.

length, elongation %, fiber strength and uniformity ratio. Moderate heritability was observed for the ginning outturn, while low heritability was documented for single plant yield. For a selection program to be more effective, it is crucial to consider both heritability estimates and predicted genetic advances. High heritability along with GAM was recorded in the number of monopodia per plant, the number of sympodia per plant, fiber fineness, boll weight, plant height, seed index and lint index. High heritability with moderate GAM was documented in days to 50 % flowering, fiber strength and upper half mean length. High heritability combined with low GAM was observed in the elongation % and uniformity ratio. Medium heritability in addition to medium GAM was noted in the ginning outturn. Low heritability and low GAM were recorded in single plant yield.

Correlation studies

The correlation coefficient is a tool used in plant breeding to ascertain the extent of the relationship between 2 or more variables. Days to 50 % flowering exhibited a significant negative correlation with single plant yield, while a substantial inverse relationship was observed between the lint index and seed index. Moreover, a significant negative correlation was noted between the number of monopodia per plant and the number of sympodia per plant. The uniformity ratio and seed index displayed significant negative correlations with the ginning outturn. Additionally, the lint index exhibited a significant positive correlation with both seed index and fiber fineness. Fiber quality traits like upper half mean length displayed positive and significant correlations with fiber strength, elongation % and uniformity ratio. Conversely, fiber fineness showed a significant negative correlation with fiber strength, upper half mean length and elongation ratio, while demonstrating a significant positive association with the lint index. Moreover, the uniformity ratio exhibited a significant positive correlation with both fiber strength and upper-half mean length. Notably, for fiber strength, elongation % displayed a positive and significant correlation with both fiber strength and upper half mean length (Fig. 2)(Table 6).

Multivariate analysis

Principal component analysis

Out of the 14 PCs, the first 5 had eigenvalues greater than one (>1) and had the maximum contribution to total variability. The corresponding eigenvalues of PC 1, 2, 3, 4 and 5 were 3.02, 1.99, 1.50, 1.31 and 1.13 respectively, which contributed 21.6 %, 14.3 %, 10.8 %, 9.4 % and 8.1 % variance to the total variation of 64.2 %.

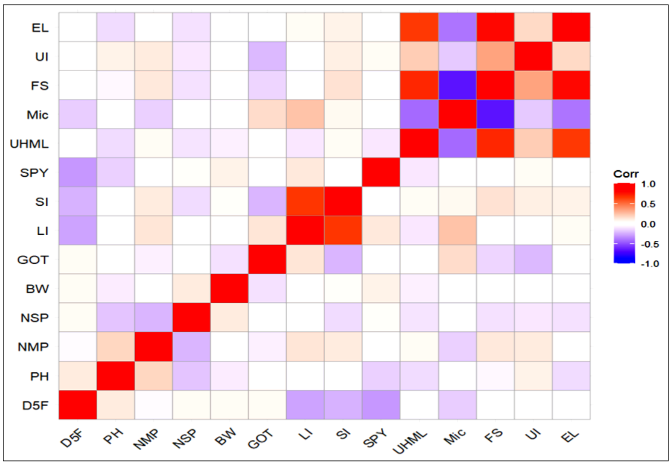

Fig. 2. Correlation heatmap for the characters studied.

Biplot

The biplot represents variables in the form of vectors on the plot. The biplot is formulated using the first 2 principal components, PC 1 and PC 2. The length of the vector from the origin depicts the contribution of each variable to the total variability, *i.e.,* the ability to discriminate the genotypes. The traits plant height, number of sympodia per plant, boll weight, single plant yield, number of monopodia per plant and uniformity ratio displayed minimum differences as they are close to the origin, whereas traits such as fiber fineness, lint index, seed index, elongation %, fiber strength, upper half mean length and days to 50 % flowering exhibited maximum differences as they are away from the origin. The genotypes SIA-5, CBS 202, ERB 4488, EC 13758, EGYPTIAN, EC 101784, EC 97626 and EC 136453 showed maximum divergence (Fig. 4).

Table 6. Variability components of the characters studied.

Traits	Mean	PV	GV	GCV (%)	PCV (%)	h^2	GA	GA as % of mean
Days to 50 % flowering	70.97	43.65	42.81	9.22	9.31	98.09	13.37	18.84
Plant height (cm)	74.29	218.44	208.83	19.45	19.89	95.6	29.15	39.23
Number of Monopodia/plants	1.89	0.51	0.47	36.5	37.72	93.61	1.38	72.85
Number of sympodia/plant	15.94	29.7	27.99	33.19	34.19	94.24	10.59	66.48
Boll weight (g)	3.41	0.54	0.51	20.96	21.66	93.64	1.42	41.84
Ginning outturn (%)	35.41	5.67	4.24	5.81	11.58	58.71	4.62	14.03
Lint index	4.97	0.33	0.3	11.08	11.54	92.16	1.09	21.95
Seed index	9.02	1.36	1.35	12.89	12.94	99.27	2.39	26.5
Single plant yield (g)	55.03	190.6	36.05	10.91	56.41	18.97	5.39	9.8
Upper half mean length (mm)	30.5	3.96	3.46	6.1	6.52	87.42	3.59	11.76
Fibre fineness (µg/inch)	3.37	0.52	0.49	20.83	21.3	95.69	1.42	42.04
Fibre strength (g/tex)	30.87	15.94	9.95	10.21	12.92	62.41	5.14	16.64
Uniformity ratio (%)	80.18	2.41	1.48	1.52	1.94	61.23	1.96	2.45
Elongation (%)	5.71	0.06	0.05	3.95	4.37	81.91	0.42	7.38

The germplasm with a high PC 1 value was characterized by fiber fineness, ginning outturn and the number of sympodia per plant, significantly contributing to the variation observed among the genotypes under investigation. PC 2 revealed higher and positive values for days to 50 % flowering, plant height, number of sympodia per plant and ginning outturn. PC 3 has high and positive values for traits: days to 50 % flowering, plant height, number of monopodia per plant, seed index, fiber fineness and uniformity ratio. PC 4 showed high and positive values for traits including ginning outturn, lint index, upper half mean length, fiber strength, fiber fineness and elongation %. PC 5 displayed high and positive values for plant height, number of monopodia per plant, ginning outturn, single plant yield, upper half mean length, fiber fineness and uniformity ratio.

Scree plot

The association between the eigenvalues and principal components is shown by a scree plot, and PC 1 had the highest variability of 21.6 % with an eigenvalue of 3.02. PC 13 and PC 14 showed the least variability, with eigenvalues of 0.18 and 0.122 respectively (Fig. 3).

Cluster analysis

A dendrogram based on Euclidean distance was constructed to separate 100 germplasm into 5 main clusters (Fig. 5). In the current study, multivariate hierarchical cluster analysis was used and the cluster analysis was performed using the mean data. The ward's linkage approach was used to cluster the data and Euclidean distance was used to compute intra- and inter-cluster distances. Cluster I contains 13 genotypes and this cluster was subdivided into 2 subclusters, I(a) and I(b). Cluster II consists of 32 genotypes and it also contains 2 subclusters II(a) and II(b) with 16 and 16 genotypes respectively. Cluster III consists of 17 genotypes, which are further divided into subclusters III(a) and III(b) with 6 and 11 genotypes respectively. Cluster IV consists of 28 genotypes, which were subdivided into 2 subclasses, IV(a) and IV(b), with 13 and 15 genotypes respectively. Cluster V consists of 10 genotypes, which were subdivided into 2 subclasses V(a) and V(b), with 4 and 6 genotypes respectively (Fig. 5).

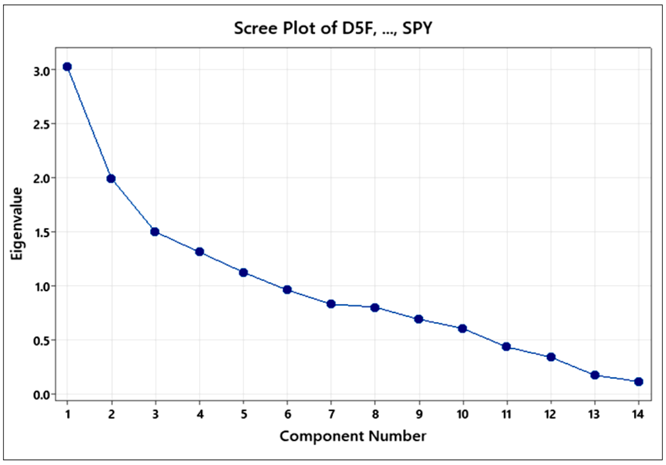

Fig. 3. Scree plot for fourteen principal components.

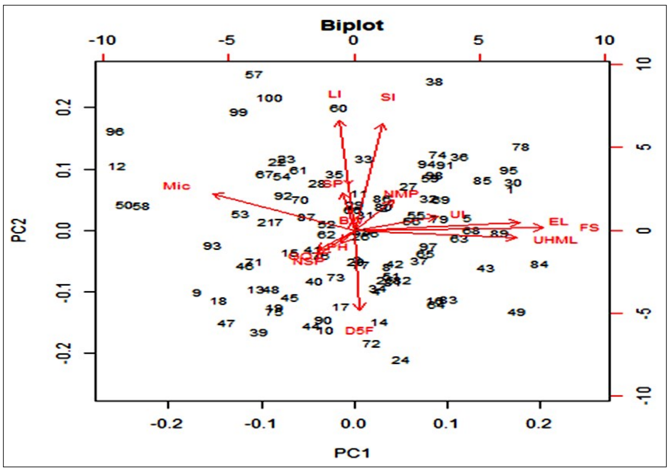

Fig. 4. Biplot between PC 1 and PC 2 displaying divergence of the traits and accessions.

Discussion

Among the accessions, the majority of genotypes had green leaves and sparse pubescence followed by medium pubescence and glabrous leaves. In general, *G. barbadense* species are highly affected by sucking pests, especially jassids and whiteflies (11). Leaf hairiness plays a prominent role in pest resistance, where the glabrous trait is associated with reduced fibre yield (12) and dense pubescence is more susceptible to whiteflies (13) and also increases gin trash. *G. barbadense* germplasm with medium pubescence can exhibit tolerance and potential genotypes can be selected for the improvement of whitefly-resistant cultivars. The majority of germplasm was identified to have a flat leaf appearance and the presence of stem and petiole pigmentation as well as petal spot indicated the usefulness of this character as a tool for the identification of specific genotypes. Among the genotypes, 60 genotypes showed semi-digitate leaves and 40 genotypes were recognized to have palmate leaf shape. Regarding boll traits, most of the germplasm accessions were observed to have elliptical and small-sized bolls. Thus, DUS descriptor characterization of germplasm is advantageous for protecting and identifying varieties.

By studying genetic variability, breeders can identify individuals or lines with desirable traits that can be used as parents for crossbreeding programs (14). Genetic variability estimates pointed out that the differences between GCV and PCV are not very wide, primarily due to genetic factors, and indicate the presence of considerable genetic variability. There is good potential for improvement through hybridization coupled with subsequent selection.

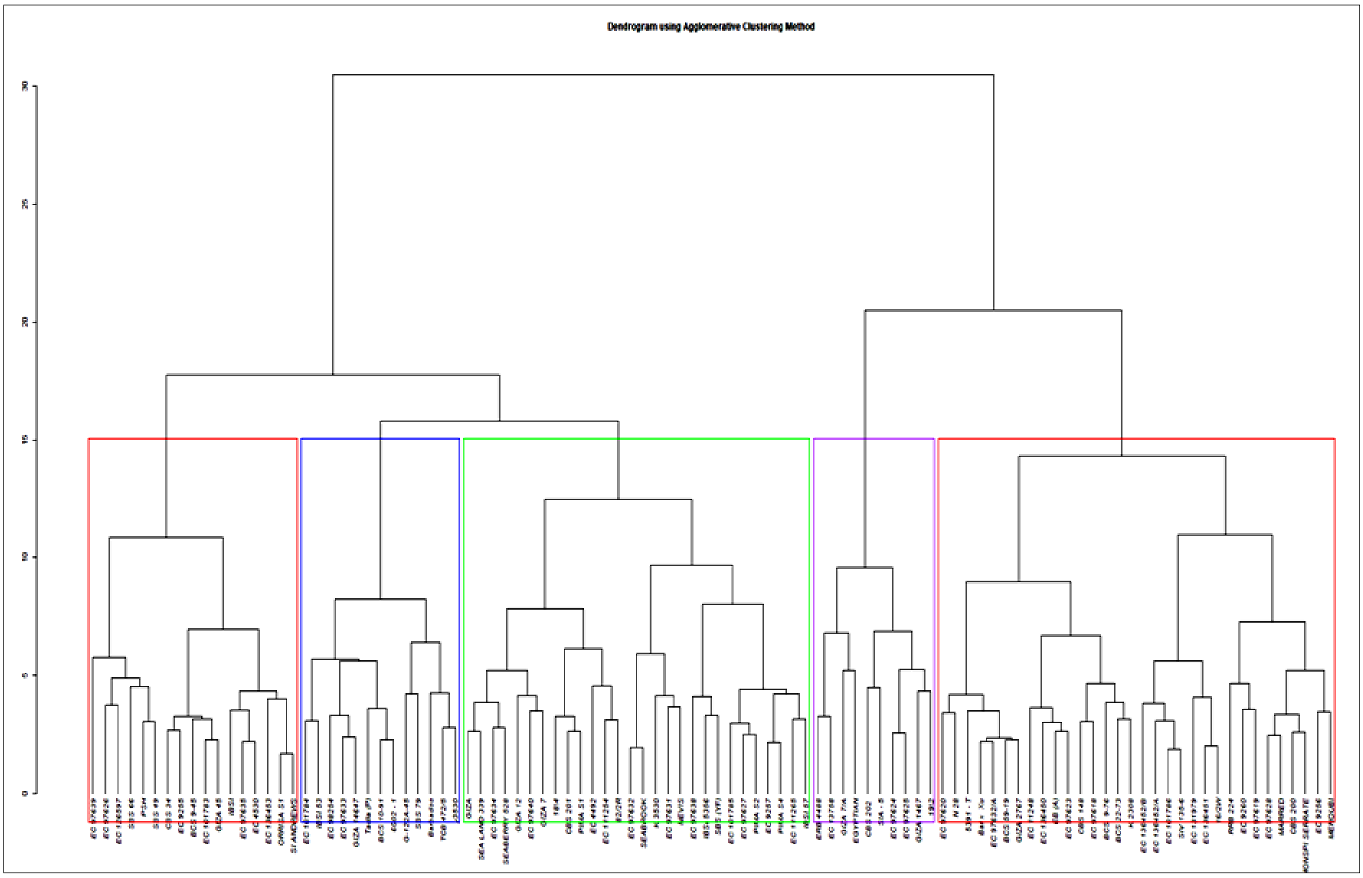

Fig. 5. Clustering of 100 *G. barbadense* germplasm accessions based on Euclidean distance.

The results also indicated that the PCV was slightly higher than the GCV for almost all traits, which indicates traits were less influenced by the environment. High GCV and PCV for traits: number of monopodia per plant, number of sympodia per plant, boll weight and fibre fineness showed an abundance of potential for genetic improvement of these traits by selection. A similar trend was observed for the number of monopodia per plant by (15-17) and the number of sympodia per plant by (18-20).

High heritability for traits like seed index, days to 50 % flowering, fibre fineness, plant height, number of sympodia per plant, boll weight, number of monopodia per plant, lint index, upper half mean length, elongation %, fibre strength and uniformity ratio were documented, which indicates the traits are less influenced by the environment and can be improved through simple selection methods. High heritability coupled with high GAM was noted in the number of monopodia per plant, the number of sympodia per plant, fibre fineness, boll weight, plant height, seed index and lint index. These findings imply, that additive gene action is more prevalent in these traits' inheritance and that environmental influence has less of an effect on the expression of these traits. Therefore, these traits are highly reliable during selection. The results were similar to (17, 21-23) for the number of monopodia/plants; (17, 22, 24) for the number of sympodia per plant; (22, 23, 25, 26) for boll weight; (21, 23, 26) for fibre fineness and (23, 24, 26) for lint and seed index.

In plant selection, computing the relationship between yield and its characteristics is crucial. Seed cotton yield per plant showed a negative and significant association with days to 50 % flowering, whereas selection for these traits results in high-yielding and early maturing genotypes as indicated by previous studies (27). The number of monopodia per plant has a significant negative correlation with the number of sympodia per plant. A significant negative correlation exists between ginning outturn and both seed index and uniformity ratio. Similar findings are reported (21, 28). The lint index has a significant positive correlation with the seed index and fibre fineness. Between seed index and lint index, a significant positive correlation was noticed. Similar findings were also reported (29). The upper half mean length significantly correlates with fiber strength, elongation % and uniformity ratio, as reported (30). The correlation between fiber fineness and lint index was significantly positive and fiber strength, upper half mean length and elongation ratio exhibited a negative correlation. A positive and significant correlation was observed between fiber strength and elongation %, upper half mean length and uniformity ratio, which was followed (31). According to correlation studies, selection focused on favourable associations between these traits would be a very effective way to increase both yields and fibre quality. Selection should be performed cautiously for those traits that have a negative correlation with seed cotton yield.

The total variance has been divided into 14 components employing PCA. Out of the 14 PCs, the 5 PCs had the highest proportion of total variability and an eigenvalue greater than one, which was similar to the results reported (32). A scree plot shows the relationship between principal components and eigenvalues; PC 1 had the greatest variability, at 21.6 %, with an eigenvalue of 3.02. PC 13 and PC 14 showed minimal variation, with eigenvalues of 0.18 and 0.122 respectively. Since PC 1 exhibited the highest variability, selection should be made using PC 1's genotypes. Maximum divergence was contributed by traits fibre fineness, lint index, seed index, elongation %, fibre strength and upper half mean length and days to 50% flowering. The genotypes 3530, SIA-5, CBS 2020, ERB 4488, EC 13758, EC 101784, EGYPTIAN, EC 97626 and EC 136453 that showed maximum divergence can be utilized to develop heterotic hybrids in future breeding programmes.

Cluster analysis showed that cluster II has the greater number of genotypes (32) while cluster V consists of the least number of genotypes (10). The maximum inter-cluster distance was observed between clusters 1 and 5 (5.12) followed by 3 and 5 (5.07). The maximum inter-cluster distance indicates that genotypes were far more diverse than other clusters. Based on genetic diversity analysis, the genotypes from clusters I and V could be utilized as parents in hybridization programmes, resulting in high heterotic combinations in subsequent generations.

Conclusion

The current study revealed genetic variability and diversity in 100 *G. barbadense* germplasm accessions. High GCV and PCV values, along with high heritability and GAM for traits such as the number of monopodia per plant, number of sympodia per plant, boll weight, and fibre fineness, suggest additive gene action plays a predominant role. This genetic makeup makes them reliable for selection purposes, with minimal influence from environmental factors. Genotypes SBS 66, SBS 79 and EC 98254 showed significantly higher values than the mean and serve as a source for the selection of parents for improving traits viz., boll weight, upper half mean length, fibre fineness and single plant yield. PCA revealed that 5 PCs had the highest proportion of total variability, of which PC 1 exhibited the highest variability. The results of cluster analysis revealed that genotypes from clusters I and V can be utilized in hybridization. Multivariate analysis reveals that substantial diversity was present in the germplasm, whereas genotypes 3530, CBS 202, ERB 4488 and EGYPTIAN can be used as parents in future breeding improvement.

Acknowledgements

I acknowledge all authors for their constant support and guidance. I am grateful to the Department of Cotton and the Department of Genetics and Plant Breeding for providing resources for the successful completion of the work.

Authors' contributions

HS collected and performed the analysis of the plant material, statistical analyses and wrote the manuscript. NP reviewed the manuscript and supervised all activities, participated in field research, data analysis and manuscript preparation. AS supervised all activities and participated in field research. NM and KG provided technical support for manuscript preparation. All authors have read and agreed to the publication of the final manuscript.

Compliance with ethical standards

Conflict of interest: Authors do not have any conflict of interests to declare.

Ethical issues: None

References

1. Hussein EHA, Osman MHA, Hussein MH, Adawy SS. Molecular characterization of cotton genotypes using PCR-based markers. J Appl Sci Res. 2007;3(10):1156-69.
2. Zada M, Zakir N, Rabbani MA, Shinwari ZK. Assessment of genetic variation in Ethiopian mustard (*Brassica carinata* A. Braun) germplasm using multivariate techniques. Pak J Bot. 2013;45 (S1):583-93.
3. Ameena N. Morphological, chemical, biochemical methods for cultivar identification in cotton (*Gossypium* sp.) [M. Sc Thesis]. [Coimbatore]: Tamil Nadu Agricultural University; 2009.
4. Dhamayanthi KPM, Manivannan A, Saravanan M. Evaluation of new germplasm of Egyptian cotton (*G. barbadense*) through multivariate genetic component analysis. Electronic Journal of Plant Breeding. 2018;9(4):1348-54. https://doi.org/10.5958/0975-928X.2018.00168.0
5. Brown-Guedira GL, Thompson JA, Nelson RL, Warburton ML. Evaluation of genetic diversity of soybean introductions and North American ancestors using RAPD and SSR markers. Crop Sci.2000;40(3):815-23. https://doi.org/10.2135/cropsci2000.403815x
6. Federer WT, Raghavarao D. On augmented designs. Biometrics. 1975;29-35. https://doi.org/10.2307/2529707
7. Protection of plant varieties and farmers' rights act (PPV & FR). Vol. 53. Delhi: Department of Agriculture and Cooperation, Ministry of Agriculture; 2001.
8. Burton GW, De Vane de EH. Estimating heritability in tall fescue (*Festuca arundinacea*) from replicated clonal material. 1953; https://doi.org/10.2134/agronj1953.00021962004500100005x
9. Johnson HW, Robinson HF, Comstock RE. Estimates of genetic and environmental variability in soybeans. 1955; https://doi.org/10.2134/agronj1955.00021962004700070009x
10. Manivannan A, Mohan P, Saravanan M, Waghmare VN. Morphological characterization of Asiatic cotton (*G. arboreum*) germplasm of India. Electronic Journal of Plant Breeding. 2018;9(2):673-81. https://doi.org/10.5958/0975-928X.2018.00081.9
11. Dhamayanthi KPM, Rameash K, Manivannan A, Sheeba A, Abirami S. Studies on leaf hairiness and sucking pest resistance in Egyptian cotton (*Gossypium barbadense* L.). 2020;
12. Bourland FM, Gbur EE. Relationships of plant trichomes to yield and fiber quality parameters in upland cotton. J Cotton Sci. 2017;21(4):296-305. https://doi.org/10.56454/FUIN8421
13. Miyazaki J, Stiller WN, Wilson LJ. Identification of host plant resistance to silverleaf whitefly in cotton: implications for breeding. Field Crops Res. 2013;154:145-52. https://doi.org/10.1016/j.fcr.2013.08.001
14. Subedi M. Developing *Chenopodium ficifolium* as a diploid model system relevant to genetic characterization and improvement of allotetraploid *C. quinoa*. 2020; https://doi.org/10.1186/s12870-021-03270-5
15. Joshi V, Patil BR. Genetic variability and heritability study in F2 population for yield, yield attributes and fibre quality traits in cotton (*Gossypium hirsutum* L.). J Pharmacogn Phytochem. 2018;7(4):2816-18. https://doi.org/10.20546/ijcmas.2018.709.044
16. Gnanasekaran M, Thiyagu K, Gunasekaran M. Genetic variability heritability and genetic advance studies in cotton (*Gossypium hirsutum* L.). Electronic Journal of Plant Breeding. 2018;9(1):377-82. https://doi.org/10.5958/0975-928X.2018.00043.1
17. Mawblei C, Premalatha N, Rajeswari S, Manivannan A. Genetic variability, correlation and path analysis of upland cotton (*Gossypium hirsutum* L.) germplasm for seed cotton yield. Electronic Journal of Plant Breeding. 2022;13(3):820-25. https://doi.org/10.37992/2022.1303.104
18. Rathinavel K. Agro-morphological characterization and genetic diversity analysis of cotton germplasm (*Gossypium hirsutum* L.). Int J Curr Microbiol Appl Sci. 2019;8(2):2039-57. https://doi.org/10.20546/ijcmas.2019.802.237
19. Saritha HS, Patil RS. Genetic studies in colour-cotton (*Gossypium hirsutum* L.) genotypes for seed cotton yield and fibre quality traits. Electronic Journal of Plant Breeding. 2020;11 (01):30-35. https://doi.org/10.37992/2020.1101.006
20. Chapara R, Reddy KV, Rani MS, Lakshmi BS, Roja V, Pranaya J. Variability studies and genetic divergence in cotton (*Gossypium hirsutum* L.) germplasm using multivariate analysis. Electronic Journal of Plant Breeding. 2022;13(4):1305-11. https://doi.org/10.37992/2022.1304.173
21. Gnanasekaran M, Thiyagu K, Gunasekaran M. Studies on genetic variability correlation and path analysis in upland cotton. Electronic Journal of Plant Breeding. 2020;11(03):981-86. https://doi.org/10.37992/2020.1103.160
22. VKI Sri Subalakhshmi SR, Premalatha N, Thirukumaran K, Boopathi NM. Genetic variability studies for yield components and fibre quality traits in upland cotton (*Gossypium hirsutum* L.). Electronic Journal of Plant Breeding. 2022;13(3):991-99. https://doi.org/10.37992/2022.1303.12823
23. Chapara R, Reddy Kvs, Rani Ms, Sudhamani K. Association and divergence studies for yield and its attributing traits in upland cotton (*Gossypium hirsutum* L.).
24. Manonmani K, Mahalingam L, Malarvizhi D, Sritharan N, Premalatha N. Genetic variability, correlation and path analysis for seed cotton yield improvement in upland cotton (*Gossypium hirsutum* L.). J Pharmacogn Phytochem. 2019;8(4):1358-61.
25. Aarthi V, Balu PA, Premalatha N. Genetic variability studies in germplasm of upland cotton (*Gossypium hirsutum* L.). Electronic Journal of Plant Breeding. 2018;9(4):1600-05. https://doi.org/10.5958/0975-928X.2018.00200.4
26. Premalatha N, Mahalingam L, Kumar M, Rajeswari S. Genetic variability studies in *Gossypium barbadense* L. genotypes for seed cotton yield and its yield components. Int J Curr Microbiol Appl Sci. 2020;9(6):1-12. https://doi.org/10.20546/ijcmas.2020.906.001
27. Kumar CPS, Prasad V, Rajan REB, Joshi JL, Thirugnanakumar S. Studies on correlation and path-coefficient analysis for seed cotton yield and its contributing traits in cotton (*Gossypium hirsutum* L.). 2019;
28. Srinivas B, Bhadru D, Rao MV. Correlation and path coefficient analysis for seed cotton yield and its components in American cotton (*Gossypium hirsutum* L.). Agricultural Science Digest-A Research Journal. 2015;35(1):13-18. https://doi.org/10.5958/0976-0547.2015.00003.8

29. Amein MMM, Masri MI, El-Adly HH, Attia SS. Correlation and path coefficient analysis for yield components traits in Egyptian cotton genotypes (*Gossypium barbadense* L.). Plant Arch. 2020;20 (2):803-06.

30. Gauswami JJ. Genetic variation, correlation, path coefficient and D2 analysis in American cotton (*Gossypium hirsutum* L.). 2020;

31. Bayyapu Reddy K. Genetic analysis of fibre quality traits in intra-specific hybrids of cotton (*Gossypium hirsutum* L.). 2015;

32. Sagar U Nikhil, Reddy BV, Vishnuvardhan K Mohan, Prasanna Rajesh A, Rao ESV, Sabitha N. Genetic analysis of fibre quality and yield related traits in Desi cotton (*Gossypium arboreum* L.). Electronic Journal of Plant Breeding. 2023;14(3):833-40. https://doi.org/10.37992/2023.1403.094