

DIVERSITY AND CORRELATION STUDIES OF INTRASPECIFIC HYBRIDS FOR THEIR MORPHOLOGICAL TRAITS IN COTTON (GOSSYPIMUM HIRSUTUM L.)

Nouman Bilal^{*1}, Wajad Nazeer², Abdul Latif Tipu³, Shoaib Liaqat⁴, Zaid Waqas⁵,
Muhammad Aurang Zaib⁶, Hafiz Usama Manzoor⁷, Aanab Fatima⁸

^{*1}Department of Plant Breeding and Genetics, Ghazi University, Dera Ghazi Khan, Pakistan

²Department of Plant Breeding and Genetics, Ghazi University, Dera Ghazi Khan, Pakistan

³Cotton Research Institute, Multan, Pakistan

⁴Cotton Research Institute, Multan, Pakistan

⁵Institute of Plant Breeding and Biotechnology, Muhammad Nawaz Sharif University of Agriculture, Multan, Pakistan

⁶Department of Plant Breeding and Genetics, Ghazi University, Dera Ghazi Khan, Pakistan

⁷Department of Plant Breeding and Genetics, Islamia University, Bahawalpur, Pakistan

⁸Department of Plant Breeding and Genetics, Ghazi University, Dera Ghazi Khan, Pakistan

^{*1}noumanbilal05@gmail.com, ²wajidpbg@yahoo.com, ³gtipu117@gmail.com,

⁴shoaibliaqat87@hotmail.com, ⁵zaidimalik14500@gmail.com, ⁶maurangzaib986@gmail.com,

⁷Sahabusama712@gmail.com, ⁸aanabrao9@gmail.com

DOI: <https://doi.org/10.5281/zenodo.21847274>

Keywords

Genetic diversity, Correlation;
Intraspecific hybrids.

Article History

Received: 01 February 2026

Accepted: 17 March 2026

Published: 31 March 2026

Copyright @Author

Corresponding Author: *

Nouman Bilal

Abstract

The present research aimed to assess genetic diversity in sixty-one crosses concerning yield and yield-related traits. The study also delved into analyzing the associations between these traits. A robust positive and highly significant correlation was observed between plant height and main stem nodes (0.91). Additionally, there was a highly significant positive relationship between the number of squares per plant and both plant height (0.51) and main stem nodes (0.58). The number of flowers exhibited a moderately positive correlation with the number of squares per plant (0.31). In terms of bolls, green bolls showed a high correlation with main stem nodes (0.35) and plant height (0.37), while open bolls (0.38) had a significant impact. Conversely, Cotton leaf curl virus disease exhibited negative correlations with other morphological traits. Genetic diversity among parents forms the foundation of any breeding program and is pivotal for increasing yield. Principal component analysis unveiled a total variance of 53% among crosses, with PC-I contributing 34.3% and PC-II contributing 18.7% of yield-related traits. The number of squares per plant, the number of green bolls, and the number of flowers showed positive significant component loading on the first two factors. Clusters I and II exhibited maximum values for the number of squares, the number of green bolls, the number of flowers, and resistance to cotton leaf curl virus disease. Crosses in cluster II displayed reasonable values for the number of open bolls and the number of flowers but showed susceptibility to cotton leaf curl virus disease. This comprehensive information, derived through the

application of various statistical tools, essentially for assessment of diversity in crosses successfully developing an improved breeding program.

Introduction

Cotton is the largest source of natural fiber and the sixth-largest source of oil, and it plays a crucial role in Pakistan's economy by contributing to employment opportunities in rural areas. Unfortunately, in recent years, Pakistan has witnessed a notable decline in cotton production, with a 41% decrease from 8.33 million bales to 4.91 million bales in last year (Pakistan Economic Survey 2022-23).

Cotton, a member of the *Gossypium* genus, includes five allotetraploid and 45 diploid species. Notably, *Gossypium hirsutum* L., also known as upland cotton, dominates global cotton cultivation, accounting for around 90% of production (Wendall *et al.*, 1992; Chen *et al.*, 2007). However, continuous breeding of *Gossypium hirsutum* L. has resulted in a narrowing of its genetic base (Seyoum *et al.*, 2018; Abdurakhmonov *et al.*, 2008; Tyagi *et al.*, 2014; Li *et al.*, 2023).

This narrowing genetic base contributes to issues such as low production and poor fiber quality (Malik *et al.*, 2014), making cotton crops vulnerable to the adverse effects of climate change (Nazir *et al.*, 2013). Recognizing the importance of genetic diversity in helping crops adapt to changing climates, there has been a recent trend towards broadening the genetic base. In this context, researchers have conducted studies involving the crossing of genetically diverse parent plants to enhance the genetic diversity of cotton, particularly in modern cultivars (Zhu *et al.*, 2019; Abdel-monaem *et al.*, 2022; Shakeel *et al.*, 2015). Assessing the diversity within germplasm is a critical step in initiating successful breeding programs (Campbell, 2009). Hybridization, the process of crossing genetically diverse parents, is a key strategy for exploiting genetic diversity, resulting in improved yields and fiber quality (Ali & Khan, 2007). Phenotypical and morphological traits serve as essential criteria for selecting distinctive parents for crossbreeding (Meredith & Bridge, 1972).

In classical statistics, the assessment of genetic variation typically involves univariate or multivariate analysis. Principal component analysis (PCA) has gained recent popularity among breeders for its ability to group genotypes and identify key contributors to genetic diversity. PCA aids in evaluating genetic diversity and identifying promising genotypes for future breeding programs. Additionally, correlation analysis is applied to morphological traits to investigate causal relationships between them (Dewey & Lu, 1959), while cluster analysis provides insights into genetic distances between genotypes, allowing for their classification into distinct groups.

In the analysis of agronomic data, both PCA and clustering techniques were employed, resulting in graphical representations of F_1 performance and their interrelationships based on agronomic parameters. This analysis provides a comprehensive overview of the genetic variability within the cotton germplasm, offering valuable insights for future breeding efforts.

Material and Methods

Experimental Site

The study was carried out at the Cotton Research Institute, Multan, during the 2023 cotton growing season. The F_1 hybrids were sown in plots of size $10' \times 2.5'$ comprising ten beds, using the boll-row sowing method. All recommended agronomic practices were followed throughout the season. Data were recorded manually on a fortnightly basis from randomly selected, tagged plants.

Trait Measurement and Statistical Analysis

Data were recorded on the following traits: plant height, first square on node number, main stem nodes, number of squares, number of flowers, number of green bolls, number of open bolls, and CLCuD (%), on a fortnightly basis. The mean of three months of data was used for statistical analysis, i.e., the cluster analysis calculated according to (Anderson & Boyes, 1957) and other most important elements estimates (PCA) through the statistical procedures by Sokal & Sneath,

(1973). The analysis was performed in r-programming.

CLCUV Disease Incidence (%) Methodology:

Disease incidence was calculated from fortnightly data using the formula given by Nazir et al. (2013).

$$DI\% = \frac{\text{Sum of all disease rating}}{\text{Total Number of plants}} \times 16.66$$

Table 1. List of intraspecific cotton hybrids developed from different parental combinations

Sr. No.	Intraspecific Hybrids	Sr. No.	Intraspecific Hybrids
1	CRP-212 × GT/Bt Okra	31	MNH-1045 × HLS
2	CRP-213 × GT/Bt Okra	32	(MNH-1090 × MNH-875) × 875
3	MNH-886/24 × Mac-07	33	(MNH-1090 × MNH-875/4) × 875/4
4	CRP-254 × GT/Bt Okra	34	CRP-212 × Super Gold
5	Super Okra × GRC-823	35	CRP-254 × CKC-6
6	Super Okra × CRI-213	36	CRP-255 × CKC-6
7	CRI-213 × Super Gold	37	(MNH-872 × SS-32) × 872
8	MNH-1050(V2 SD) × MNH-Super gold	38	CRP-212 × FH-333
9	CRI-823 × MNH-Super gold	39	MNH-872 × Mac-07
10	CRI-134 × MNH-Super gold	40	CRP-212 × CKC-6
11	CRI-212 × MNH-875	41	CRP-213 × CKC-6
12	Super Gold × CRI-212	42	CRP-213 × Super Gold
13	Super Gold × CRI-276	43	CRP-214 × Super Gold
14	Super Gold × CRI-875/6	44	CRP-254 × Super Gold
15	GRC-823 × CRI-875/6	45	CRP-255 × Super Gold
16	Super Okra × CRI-276	46	CRP-212 × FH-333
17	BC5F1(211009) × BtGT-8/23	47	(MNH-872 × SS-102) × 872
18	(MNH-1027 × CRI-276) × VR-5	48	(MNH-872 × MNH-875/4) × 875/4
19	MNH-1040 (valvet) × VR-5	49	MNH-Super Gold × CRP-208
20	IS-29 × VR-5	50	MNH-Super Gold × CRP-2012
21	217007 × VR-5	51	MNH-Super Gold × CRP-213
22	IS-12 × VR-5	52	MNH-Super Gold × CRP-214
23	CRI-234 × Super Gold	53	MNH-Super Gold × CRP-254
24	CRI-276 × Super Gold	54	CRI-269 × MNH-Super gold
25	MNH-1045 × VR-5	55	CRI-276 × MNH-Super gold
26	MNH-1045 × VR-16	56	CRI-287 × CRI-212
27	HLS × GRC-823	57	CRI-288 × CRI-294
28	1045 R × CRI-276	58	CRI-289 × CRI-213
29	Super GOLD × 1045 R	59	CRI-290 × CRI-291
30	Super GOLD × HLS	60	CRI-294 × CRI-212
		61	CRI-287 × CRI-212

Result and Discussion

Mean Performance of Intraspecific Hybrids

The tallest cross observed was (MNH-872 × MNH-875/4) × 875/4, measuring (95.0 cm) in height, followed closely by (MNH-1090 × MNH-875) × 875, which stood at 90.2 cm (Fig 1). In contrast, crosses with shorter heights included MNH-Super Gold × CRP-212, 50.1 cm and CRI-276 × Super Gold, 55.8 cm. The mean values for first square on node number within crosses, with the highest node numbers found in Super Okra × CRI-213 (10.33), CRP-212 × GT/Bt Okra (10.00), and Super Gold × CRI-875/6 (10.00) (Fig 2). Conversely, crosses such as CRP-254 × Super Gold, CRP-255 × Super Gold, CRP-212 × FH-333, and (MNH-872 × SS-102) × 872 exhibited the lowest node numbers where the first sympodial/squares formed, measuring (6.00).

For main stem nodes, (MNH-872 × MNH-875/4) × 875/4 displayed the highest number with (30.00), followed by (MNH-1090 × MNH-875) × 875 with (27.17) (Fig 3). On the other hand, Super Gold × CRI-212 and MNH-Super Gold × CRP-2012 had the minimum numbers of main stem nodes, measuring (18.17) and (15.75), respectively. Regarding the number of squares, the highest mean values were observed in (MNH-1090 × MNH-875) × 875 (23.50), Super Gold × CRI-875/6 (22.58), MNH-886/24 × Mac-07 (21.67), 217007 × VR-5 (21.49), IS-12 × VR-5 (21.21), and

(MNH-1090 × MNH-875/4) × 875/4 (19.830). In contrast, crosses like CRP-212 × CKC-6 (11.22) and Super Okra × GRC-823 (10.83) exhibited the lowest mean values in terms of squares (Fig. 4).

When it came to the number of flowers, CRI-269 × MNH-Super Gold displayed the highest mean value with 3.52, followed by Super Okra × GRC-823 with 3.33. Conversely, crosses such as CRI-287 × CRI-212 and MNH-Super Gold × CRP-212 exhibited lower mean values, measuring 0.57 and 0.33, respectively (Fig. 5).

In terms of green bolls, the highest mean value was noted in (MNH-872 × MNH-875/4) × 875/4 with 10.17, followed by CRP-255 × Super Gold with 8.83 (Fig 6). Conversely, CRP-254 × CKC-6 had the lowest number of bolls with a mean value of 1.28, and MNH-1040 (valvet) × VR-5 had 0.83 bolls on average. High numbers of open bolls exhibit by CRI-269 × MNH-Super gold (5.29), followed by CRP-213 × Super Gold (5.25) lowest number of open bolls mean value found in (Fig 7) CRP-254 × GT/Bt Okra (0.00), Super Okra × GRC-823 (0.00).

Lastly, the percentage of Cotton leaf curl virus disease (CLCuD%), with GRC-823 × CRI-875/6 exhibiting the highest mean value at 35.40%, followed by CRP-213 × GT/Bt Okra (33.32%) and Super Okra × CRI-276 (33.32%). Conversely, the cross CRI-289 × CRI-213 exhibited the lowest disease incidence (Fig. 8).

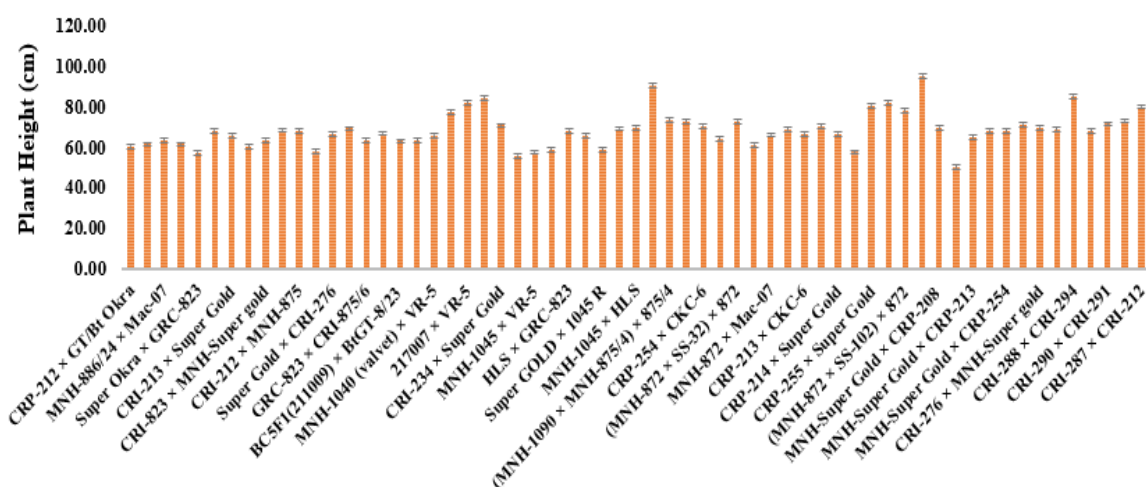


Figure 1. Mean plant height (cm) in intraspecific hybrids

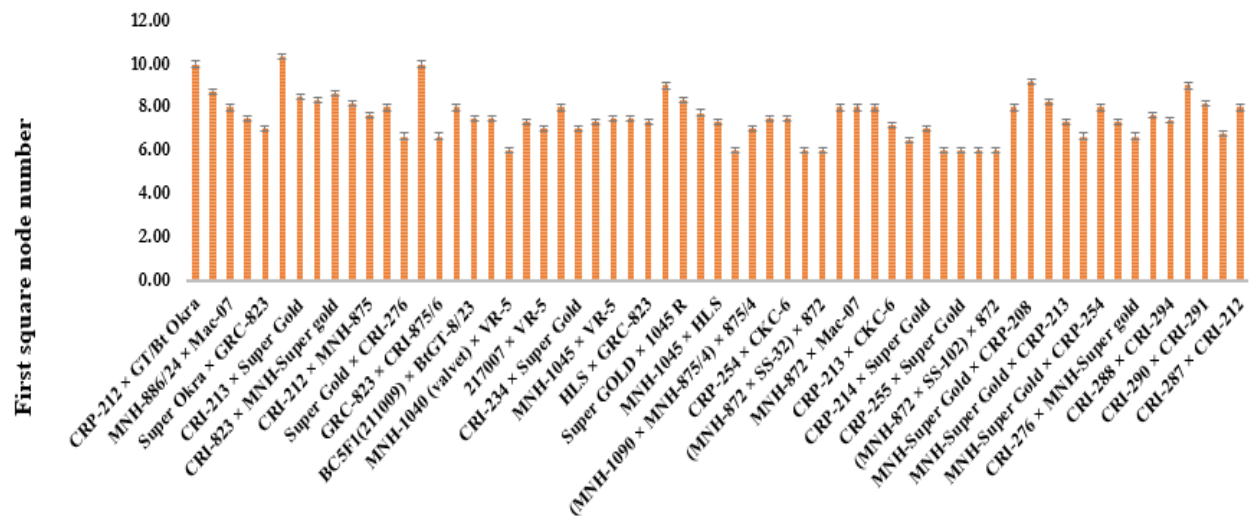


Figure 2. Mean first square on node number in intraspecific hybrids

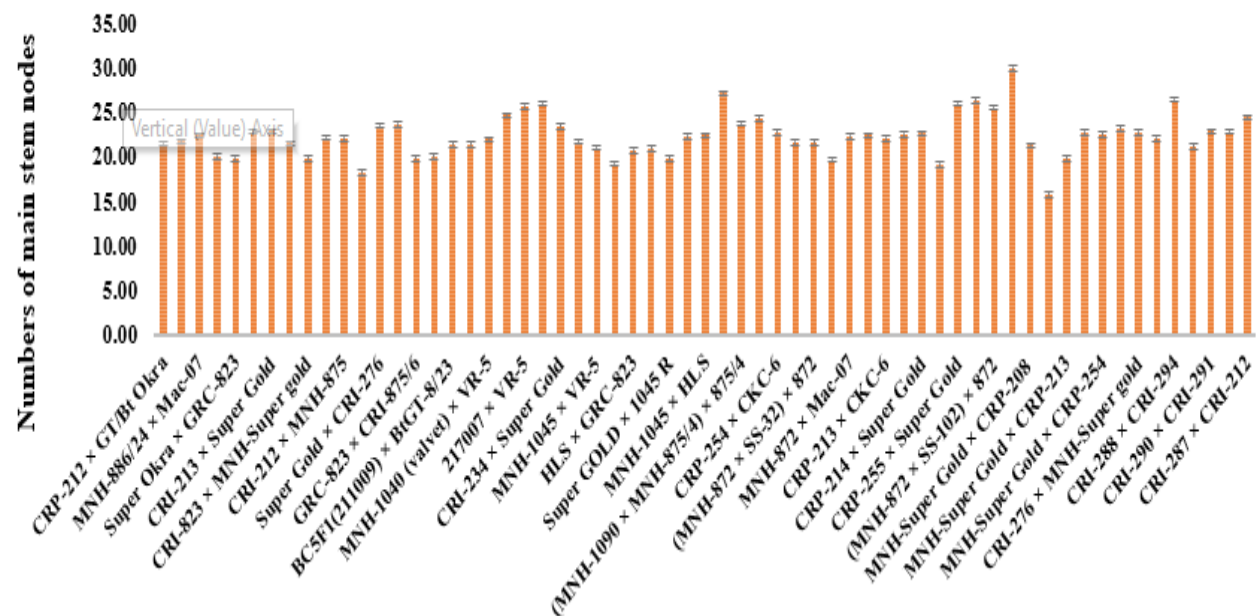


Figure 3. Mean main stem nodes in intraspecific hybrids

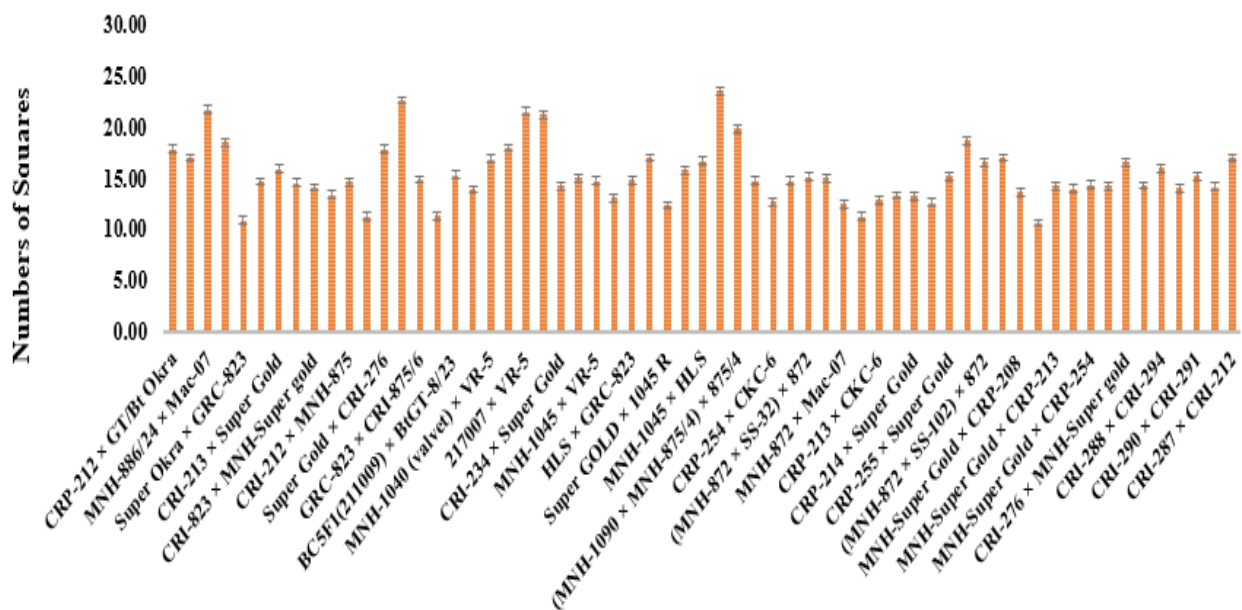


Figure 4. Mean numbers of squares in intraspecific hybrids

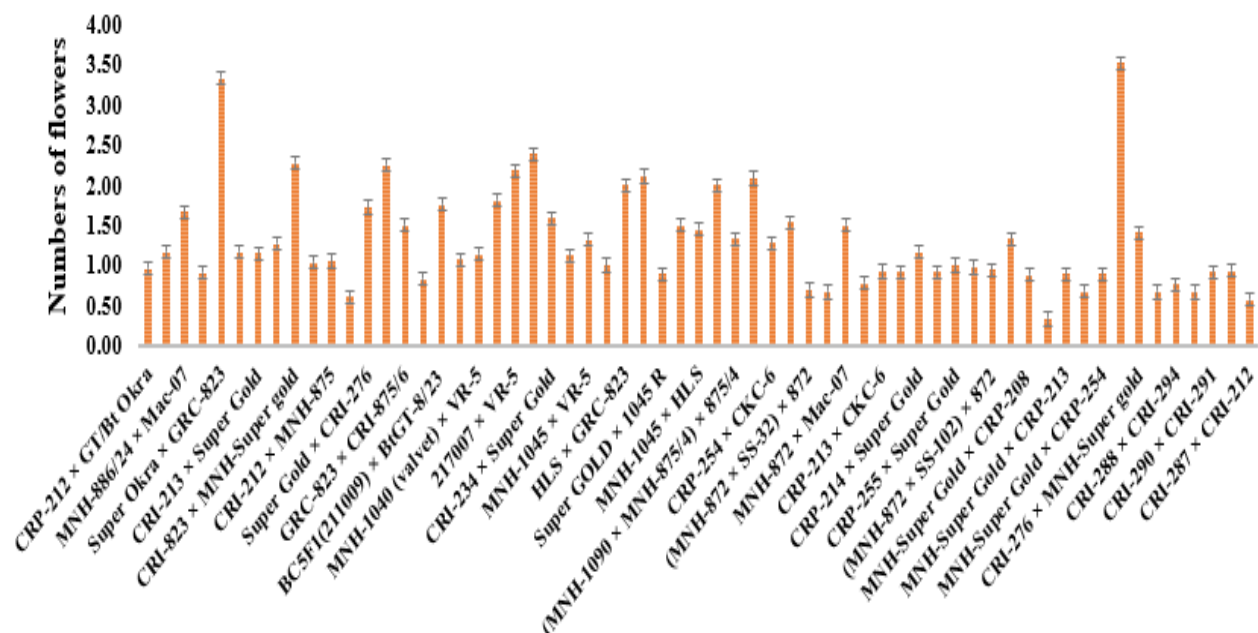


Figure 5. Mean numbers of flowers in intraspecific hybrids

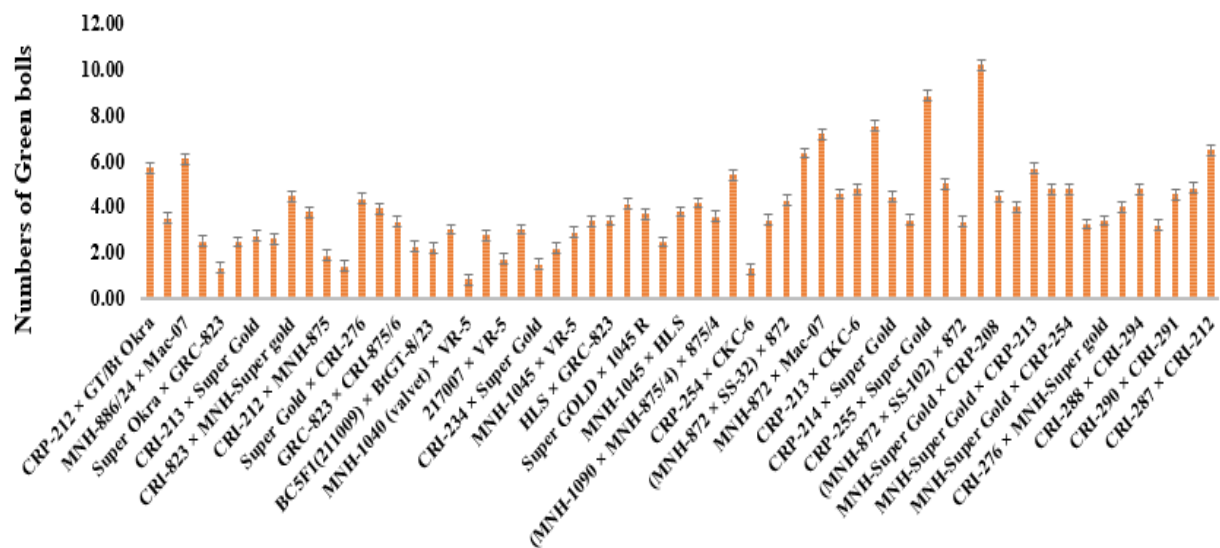


Figure 6. Mean number of green bolls in intraspecific hybrids

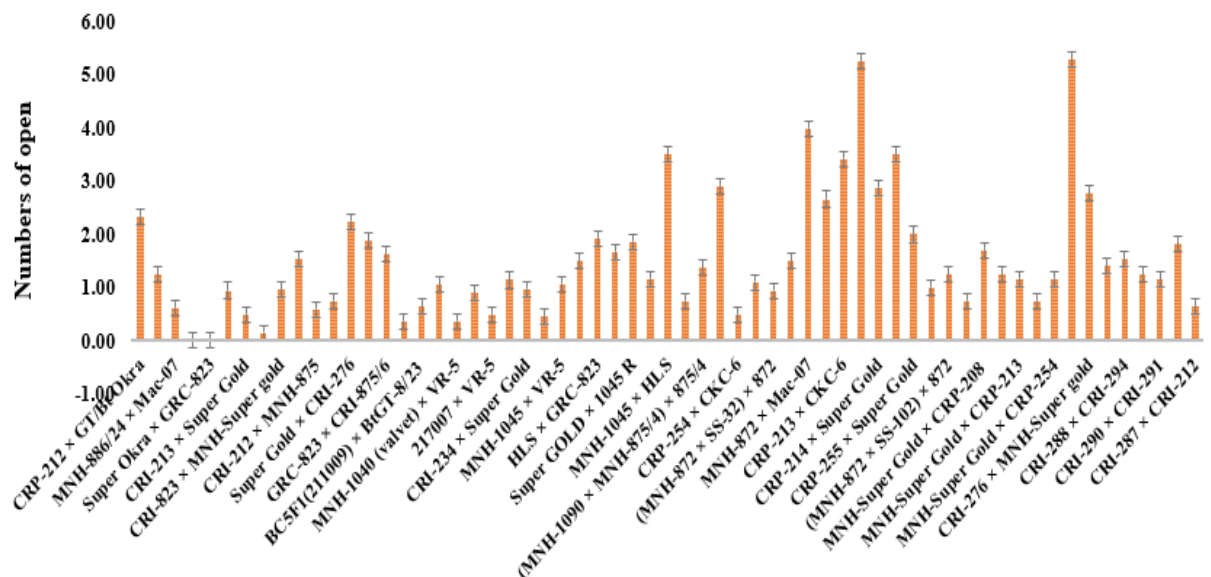


Figure 7. Mean numbers of open bolls in intraspecific hybrids

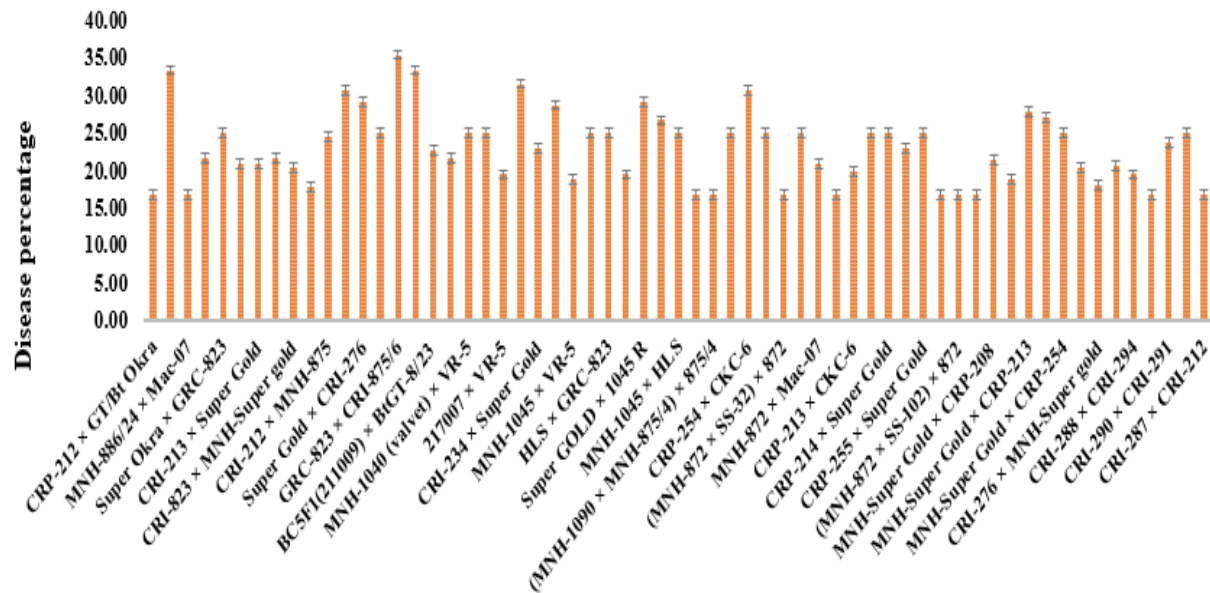


Figure 8. Mean Cotton leaf curl virus disease % in intraspecific hybrids

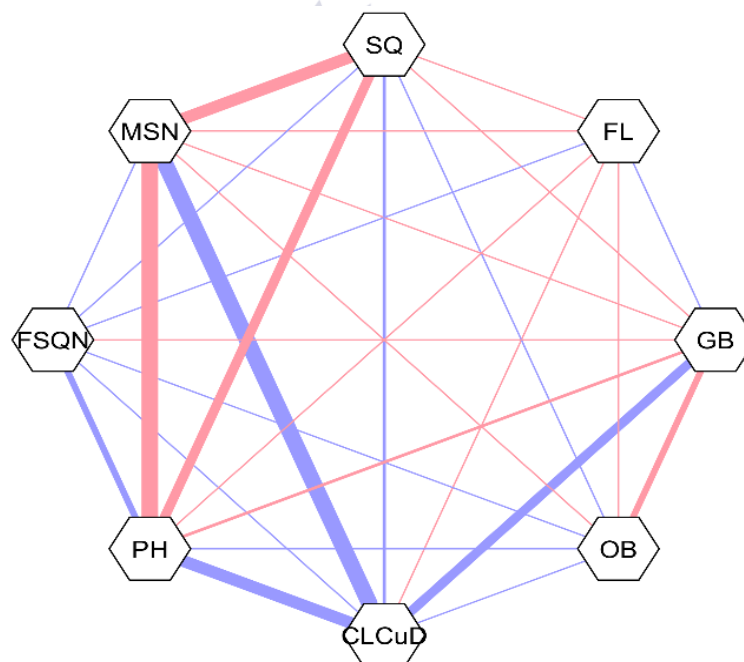


Figure 9. Pearson's correlation network showing the relationships among morphological traits in intraspecific cotton hybrids

Correlation Network Based on Pearson correlation coefficient. Showing the correlation between agronomical traits. The Pink one showing positive correlation blue one showing negative correlation and thick the line between them strong the correlation thin lines showing weak correlation.

Pearson Correlation Analysis

Pearson correlation was applied to eight agronomic traits across 61 cotton crosses, which showed sufficient variability. The results of the simple correlation analysis align closely with those obtained from PCA (Table 2). Positive correlations are illustrated in pink and negative correlations in blue (Fig. 9). The strength of the correlation between traits is indicated by line thickness: thicker lines represent stronger correlations and thinner lines represent weaker relationships. Plant height showed a positive and highly significant correlation with main stem nodes, consistent with the findings of Saeed *et al.*, (2015). The number of squares showed a positive and highly significant relationship with plant height and main stem nodes (Fig. 10). A weak but significant relationship was observed between the number of flowers and the number of squares, whereas the number of flowers showed a weak,

non-significant relationship with plant height and main stem nodes. The number of open bolls showed a weak, negative, and non-significant relationship with the number of squares. Cotton leaf curl virus disease showed a significant and highly negative relationship with plant height, main stem nodes, and green bolls. CLCuD% also showed a weak, negative, and non-significant relationship with the number of open bolls and the number of squares; similar results were reported by (Sarwar *et al.*, 2021) and (Nazir *et al.*, 2013). The first square on node number showed a negative but weak association with the other traits. The Pearson correlation coefficient was applied across seven traits to compare their relationships. The correlation matrix was calculated using the R package 'metan', and the correlation network was generated using the MetScape app in Cytoscape.

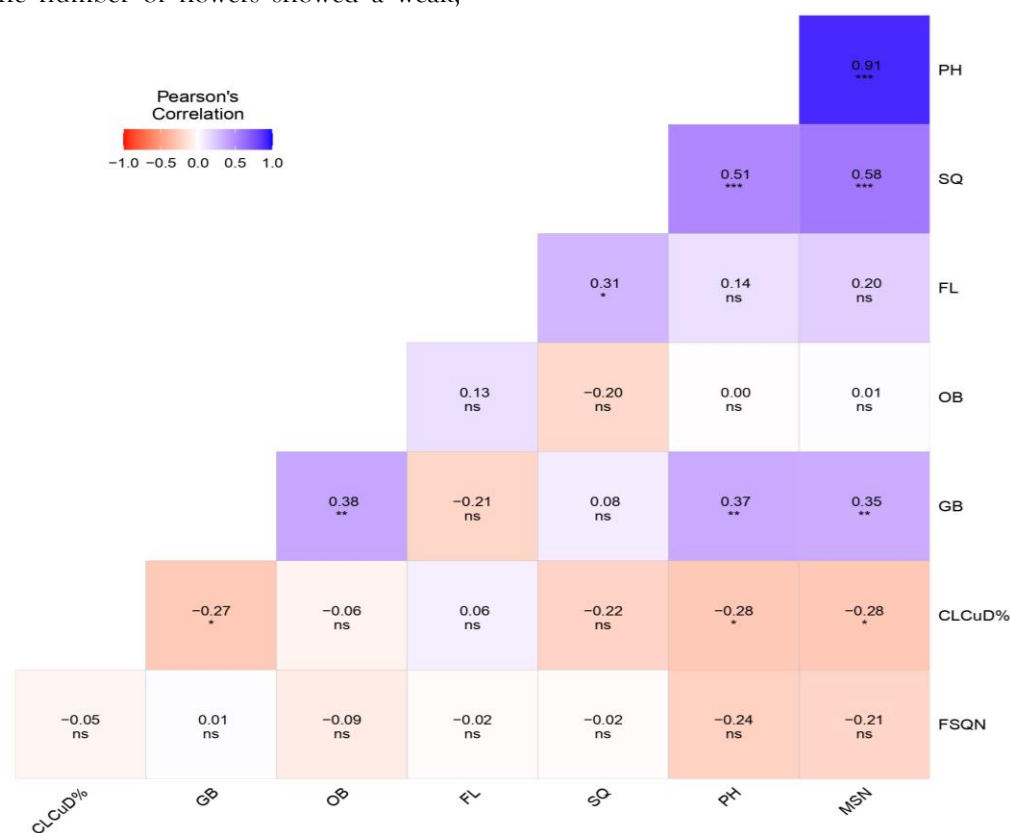


Figure 10. Pearson's correlation matrix showing the relationships among morphological traits in intraspecific cotton hybrids

***, **, * significant level 0.1%, 1% and 5% respectively, ns= nonsignificant

Principal Component Analysis to Assess the Diversity in Intraspecific Hybrids

Principal component analysis was performed on the morphological traits of cotton. PCA is a multivariate statistical approach used to examine complex, large datasets based on the associations between variables. In this study, PCA was applied to the cotton data to study the variation and identify the variables contributing to that variability.

Principal component analysis identifies groups of genotypes with the characteristics required for breeding and helps in the selection of desirable

genotypes (Imran *et al.*, 2012). PCA reduces the observations to a manageable number of components. Components with lower variance contribute less to explaining the variability among the variables than those with higher variance. Therefore, components with an eigenvalue of less than 1 are considered trivial and are excluded from further interpretation. Eigenvalues serve as a quantitative measure of how much variance each component represents. The higher the eigenvalue of a component, the greater the proportion of variance it represents (Zhi *et al.*, 2016).

Table 2. Eigenvalues, cumulative variability, and principal component loadings for morphological traits in intraspecific cotton hybrids

	PC1	PC2	PC3	PC4
Eigenvalues	1.656	1.222	1.078	1.016
Variability (%)	0.343	0.187	0.145	0.129
Cumulative (%)	0.343	0.529	0.675	0.804
Loading factor				
PH	0.554	-0.023	0.003	-0.16
FSQN	-0.142	-0.011	-0.497	0.705
MSN	0.566	-0.062	0.018	-0.086
SQ	0.421	-0.364	-0.127	0.167
FL	0.145	-0.402	0.516	0.531
GB	0.284	0.6	-0.068	0.075
OB	0.043	0.535	0.55	0.338
CLCuD%	-0.269	-0.237	0.405	-0.2

FSQN = First square on node number; FL = Number of flowers; SQ = Number of squares; PH = Plant height; MSN = Main stem nodes

OB = Number of open bolls; GB = Number of green bolls; CLCuD (%) = Cotton leaf curl disease incidence

Scree Plot

Components with higher eigenvalues account for a greater proportion of the total variation (Zhi *et al.*, 2016). Four of the eight PCs had eigenvalues greater than 1 and together accounted for 80.83% of the total variation, while the remaining PCs contributed 19.62%. PC1 accounted for 34.4% of the variation with an eigenvalue of 1.66 (Table 2), followed by PC2 (18.7%, eigenvalue 1.22), PC3 (14.5%, eigenvalue 1.08), and PC4 (12.9%, eigenvalue 1.02), as shown in the scree plot (Fig. 12).

PCA Biplot

The characteristic plant height, main stem nodes and number of squares have positive correlation of factor loading on PC II while Cotton leaf curl virus disease % and first square on node number displayed significantly negative loading on PC I. Similarly, PC II explain most of the variance by number of green bolls and number of open bolls with positive loading on PC II while other variables have negative factor loading on PC II (Table 2).

In PC III the variables Numbers of flowers and Numbers of open bolls displayed positive factor loading while negative factor loading observed for

First square on node number, Numbers of squares and Numbers of green bolls. Similarly, PC IV was explained by the variance of Number of flowers having positive factor loading on PC IV while Plant height and Main stem nodes discovered to be negative factor loading on PC IV.

A PC biplot (Fig.11) demonstrated hybrids and variables are superb and inflict as a vector on plot. The gap within traits regarding PC I and PC II depicted contribution of these traits in creating variation of hybrids, Furthermore the length of the vector presented the proportion of variation of the specific variables. However, the overall Biplot diagram demonstrated that the Number of green bolls main stem nodes and squares significantly toward the genetic diversity within germplasm and Cotton leaf curl disease% and first square on node number varying from other variables.

The PCA was conducted to dispense the aggregate variance into PCs which helps in selection of superior germplasm which would use for breeding program. In PCA the eigen value is the factor which explain variance for specific trait. The PCA is powerful tool provide easiness for obtaining suitable genotypes to initiate breeding programme (Nazir *et al.*, 2013).

Many researchers applied this method to formulize the diversity within germplasm Our

present research also testifies the robustness of similar methods and in line with the previous findings obtained by (Saeed *et al.*, 2015; Nadeem *et al.*, 2022; Zafar *et al.*, 2021; Jarwar *et al.*, 2019).

PCA Scatter Plot

The scatter plot depicted that the hybrids which were close together are considered to be similar while examined against seven characters. Hybrids 26(MNH-1045 × VR-16), 29(Super GOLD × 1045 R), 50(MNH-Super Gold × CRP-212),12(Super Gold × CRI-212),14(Super Gold × CRI-875/6),20(IS-29 × VR-5),21(217007 × VR-5),22(IS-12 × VR-5),32((MNH-1090 × MNH-875) × 875),33((MNH-1090 × MNH-875/4) × 875/4),48((MNH-872 × MNH-875/4) × 875/4),47((MNH-872 × SS-102) × 872),45(CRP-255 × Super Gold),46(CRP-212 × FH-333),57(CRI-288 × CRI-294),61(CRI-287 × CRI-212),44(CRP-254 × Super Gold),39(MNH-872 × Mac-07),42(CRP-213 × Super Gold),41(CRP-213 × CKC-6),40(CRP-212 × CKC-6) (Table 1) were farther from origin in term of eight character significantly different from the other hybrids (Fig. 11) Furthermore, the crosses which forms group near the vector which means the variation due to the variables is high in these crosses. Similiter method was adopted by Shakeel *et al.*, (2015) and Nadeem *et al.*, (2022).

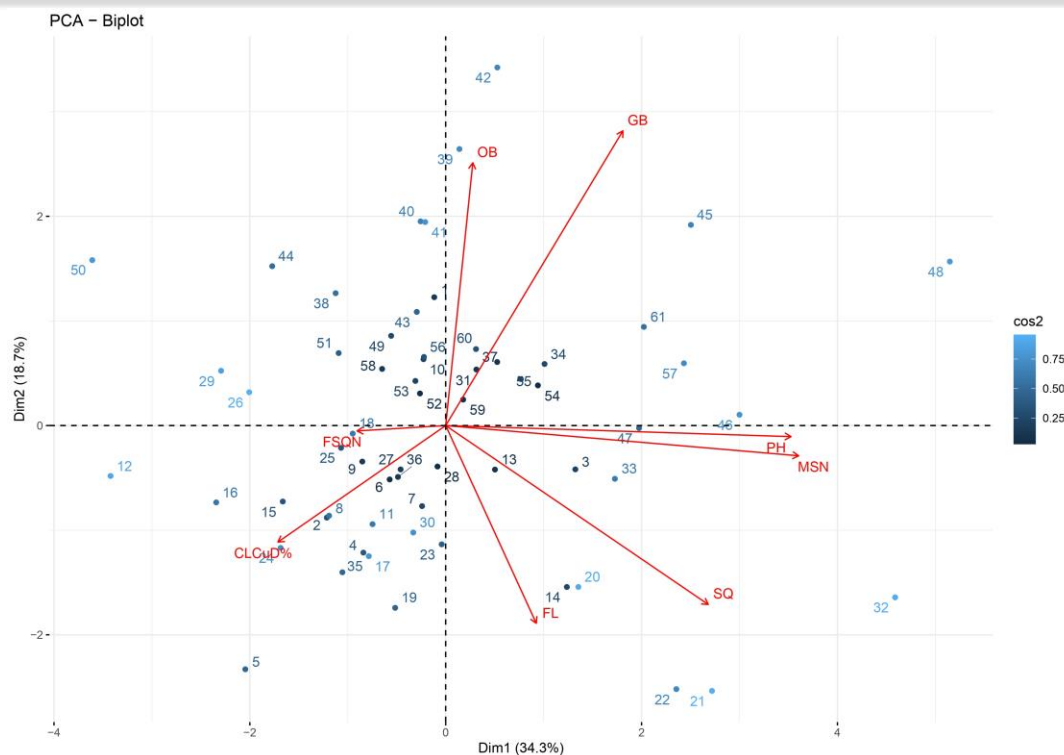


Figure 11. Biplot of PC1 and PC2 showing the contribution of morphological traits to the genetic variability of intraspecific cotton hybrid

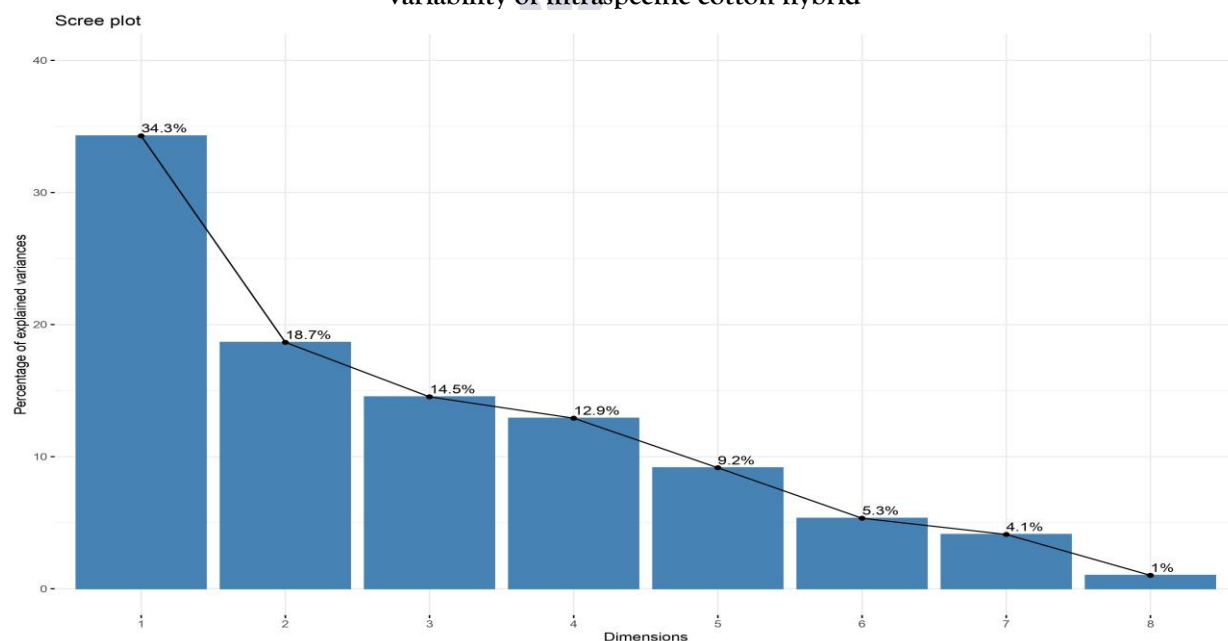


Figure 12. Scree plot showing the contribution of the first eight principal components (PCs) to the total variability in intraspecific cotton hybrids

Cluster Analysis in Intraspecific Hybrids of Cotton

Cluster analysis performed on mean values of 61 hybrids of cotton comprises from different

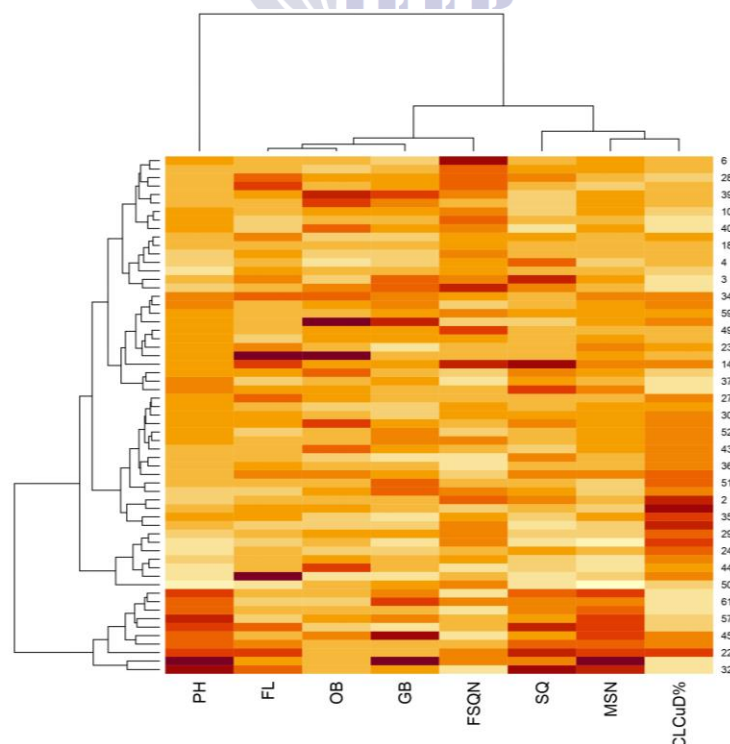
parents. The letter 1 to 61 represents the crosses (Table 1). May *et al.*, (1995) reported the cotton genotype based on close to each other grouped in same cluster. Nazir *et al.*, (2013) grouped 70

crosses and characterize them in ward's linkage method. Shakeel et al., (2015) assessed the ward linkage cluster analysis and divide the similar genotype into same group. (Nadeem *et al.*, 2022) used similar method and characterize twenty genotypes of cotton into two main cluster with division of sub cluster.

Explanation

The 61 hybrids, derived from different parents, were grouped into three clusters (Table 3). The cluster I contained 12 crosses, Cluster II comprised 11 crosses, and cluster III comprised 38 crosses. The crosses in cluster I showed reasonable values for the number of squares, number of green bolls, and number of flowers. The members of cluster I showed resistance to cotton leaf curl virus disease. The crosses in cluster II showed reasonable values for the number of open bolls and number of flowers, but these crosses also showed susceptibility to cotton leaf curl virus disease. The member of cluster III showed shorter height and fewer main stem nodes

but had high values for number of open bolls and number of flowers; some members of this cluster also had high values for first square on node number but were highly resistant to CLCuD%. Cluster III crosses can further be used for early maturity owing to their higher number of open bolls compared with the other clusters. This wide variation in germplasm is useful for selecting cotton for a range of breeding purposes. Based on cluster dendrogram (Fig. 14) these crosses are highly recommended for further use in breeding programme are IS-29 × VR-5, 217007 × VR-5, IS-12 × VR-5, (MNH-1090 × MNH-875) × 875, (MNH-1090 × MNH-875/4) × 875/4, (MNH-872 × SS 32) × 872, CRP-255 × Super Gold, CRP-212 × FH-333, (MNH-872 × SS-102) × 872, (MNH-872 × MNH-875/4) × 875/4, CRI-288 × CRI-294, CRI-287 × CRI-212. These crosses possess the quality of high flowering, high numbers of green bolls and numbers of open bolls with highly resistant to CLCuD% (Fig. 13). The members of other cluster can be used for various other traits for selection.



Tabure 13. Hierarchical cluster heatmap of intraspecific cotton hybrids showing the relatedness among morphological traits.

Table 3. Cluster membership of intraspecific cotton hybrids based on morphological traits using Ward's hierarchical clustering

Cluster I	12	IS-29 × VR-5, 217007 × VR-5, IS-12 × VR-5, (MNH-1090 × MNH-875) × 875, (MNH-1090 × MNH-875/4) × 875/4, (MNH-872 × SS-32) × 872, CRP-255 × Super Gold, CRP-212 × FH-333, (MNH-872 × SS-102) × 872, (MNH-872 × MNH-875/4) × 875/4, CRI-288 × CRI-294, CRI-287 × CRI-212.
Cluster II	11	Super Gold × CRI-276, MNH-1045 × HLS, CRP-212 × Super Gold, MNH-872 × Mac-07, CRP-212 × CKC-6, CRP-213 × CKC-6, CRP-213 × Super Gold, CRP-214 × Super Gold, CRP-254 × Super Gold, CRI-269 × MNH-Super gold, CRI-276 × MNH-Super gold.
Cluster III	38	CRP-212 × GT/Bt Okra, CRP-213 × GT/Bt Okra, MNH-886/24 × Mac-07, CRP-254 × GT/Bt Okra, Super Okra × GRC-823, Super Okra × CRI-213, CRI-213 × Super Gold, MNH-1050(V2 SD) × MNH-Super gold, CRI-823 × MNH-Super gold, CRI-134 × MNH-Super gold, CRI-212 × MNH-875, Super Gold × CRI-212, Super Gold × CRI-875/6, GRC-823 × CRI-875/6, Super Okra × CRI-276, BC5F1(211009) × BtGT-8/23, (MNH-1027 × CRI-276) × VR-5, MNH-1040 (velvet) × VR-5, CRI-234 × Super Gold, CRI-276 × Super Gold, MNH-1045 × VR-5, MNH-1045 × VR-16, HLS × GRC-823, 1045 R × CRI-276, Super GOLD × 1045 R, Super GOLD × HLS, CRP-254 × CKC-6, CRP-255 × CKC-6, CRP-212 × FH-333, MNH-Super Gold × CRP-208, MNH-Super Gold × CRP-2012, MNH-Super Gold × CRP-213, MNH-Super Gold × CRP-214, MNH-Super Gold × CRP-254, CRI-287 × CRI-212, CRI-289 × CRI-213, CRI-290 × CRI-291, CRI-294 × CRI-212.

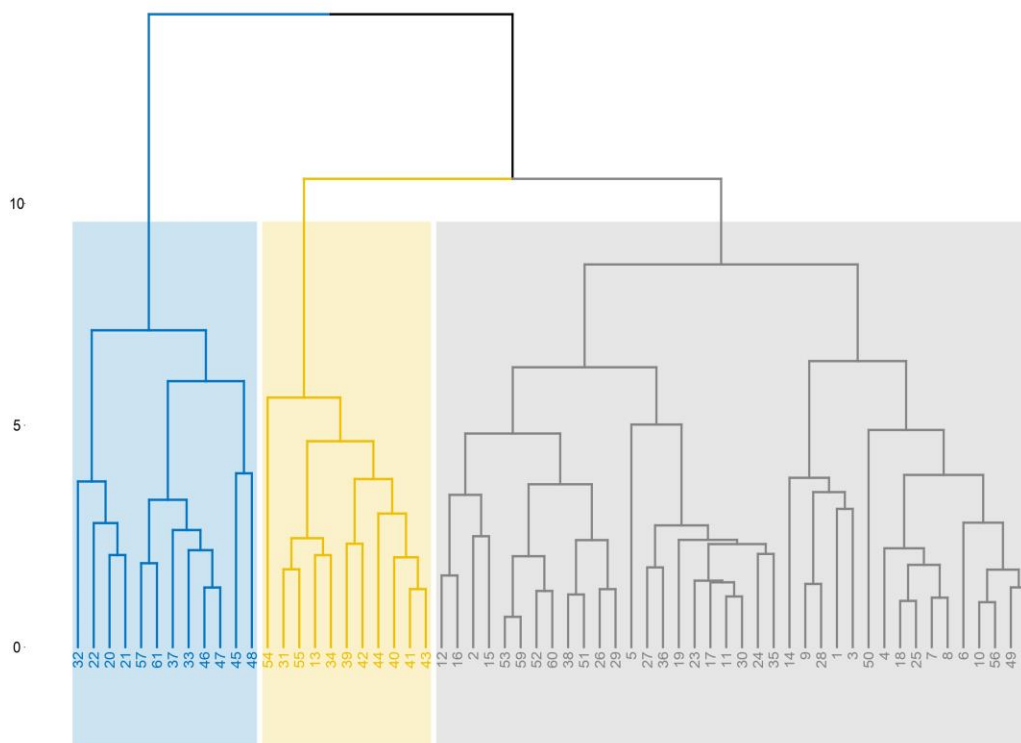
Cluster Dendrogram
15

Figure 14. Dendrogram generated using Ward's linkage method showing the clustering pattern and genetic variability among intraspecific cotton hybrids

Conclusion

In this study, we aimed to explore genetic divergence in intraspecific crosses and analyze their trait associations. Positive and significant correlations were found between plant height and main stem nodes (0.91), as well as between the number of squares per plant and both plant height (0.51) and main stem nodes (0.58). Moderate positive correlations existed between the number of flowers and squares per plant (0.31). Green bolls showed high correlations with main stem nodes (0.35) and plant height (0.37), whereas open bolls (0.38) also had a significant impact. Conversely, Cotton leaf curl virus disease exhibited negative correlations with other morphological traits.

Genetic diversity is crucial for breeding programs and yield improvement. Principal component analysis revealed 53% total variance among PC- I and PC- II, with PC-I contributing 34.3% and PC- II contributing 18.7% of yield-related traits. Squares per plant, green bolls, and flowers showed significant positive loading on these factors.

Clusters I and II demonstrated high values for squares, green bolls, flowers, and resistance to cotton leaf curl virus disease. Cluster II, while having reasonable open boll and flower numbers, was susceptible to the cotton leaf curl virus.

FUNDING: Nil

ACKNOWLEDGMENTS

We acknowledge Cotton Research Institute Multan for their technical support to conduct experiment and support during the whole experiment.

AUTHORS CONTRIBUTION

Bilal N, Nazeer W, Tipu AL, and Liaqat S conceived and designed the study, conducted the field experiment, collected and analyzed the data. Bilal N prepared the first draft of the manuscript. Nazeer W supervised the research and carried out the final editing of the manuscript. Zaib MA participated in field data collection and assisted with manuscript editing, language and spelling checks, and manuscript revision to improve originality. Waqas Z, Manzoor HU, and Fatima A

edited the manuscript, performed language and spelling checks, and improved its overall presentation. Final approval of the manuscript for publication was provided by Nazeer W. All authors read and approved the final manuscript.

CONFLICT OF INTEREST

Authors have no conflict of interest for this manuscript

REFERENCES

- Abdel-Monaem, M. A., Abido, W. A. E., Hadházy, Á., Ghoneima, M. H., EL-Mansy, Y. M., & EL-Shazly, M. W. (2022). Genetic divergence among Egyptian cotton genotypes under water deficit conditions. *Acta Ecologica Sinica*, 42(2), 11-18.
- Abdurakhmonov IY, Kohel RJ, Yu JZ, et al. Molecular diversity, and association mapping of fibre quality traits in exotic *G. hirsutum* L. germplasm. *Genomics*. 2008; 92:478-87.
<https://doi.org/10.1016/j.ygeno.2008.07.013>.
- Ali, M.A., and I.A. Khan. 2007. Assessment of genetic variation and inheritance mode in some metric traits of cotton (*Gossypium hirsutum* L.). *J. Agric. Soc. Sci.* 4(3):112-116.
- Anderson, E. and J.W. Boyes. 1957. An experimental investigation of judgments concerning genera and species. *Evolution*, 11(2): 260-263.
- Campbell, B. T., Williams, V. E., & Park, W. (2009). Using molecular markers and field performance data to characterize the Pee Dee cotton germplasm resources. *Euphytica*, 169(3), 285-301.
<https://doi.org/10.1007/s10681-009-9917-4>
- Chen ZJ, Scheffler BE, Dennis E, et al. Toward sequencing cotton (*Gossypium*) genomes. *Plant Physiol.* 2007; 145:1303-10.
<https://doi.org/10.1104/pp.107.107672>

- Chen, Z. J., Scheffler, B. E., Dennis, E., Triplett, B. A., Zhang, T., Guo, W., Chen, X., Stelly, D. M., Rabinowicz, P. D., Town, C. D., Arioli, T., Brubaker, C., Cantrell, R. G., Lacape, J.-M., Ulloa, M., Chee, P., Gingle, A. R., Haigler, C. H., Percy, R., ... Paterson, A. H. (2007). Toward Sequencing Cotton (*Gossypium*) Genomes: Figure 1. *Plant Physiology*, 145(4), 1303-1310. <https://doi.org/10.1104/pp.107.107672>
- Dewey, D. R., & Lu, K. H. (1959). A Correlation and Path-Coefficient Analysis of Components of Crested Wheatgrass Seed Production 1. *Agronomy Journal*, 51(9), 515-518. <https://doi.org/10.2134/agronj1959.00021962005100090002x>
- Dewey, D. R., & Lu, K. H. (1959). A Correlation and Path-Coefficient Analysis of Components of Crested Wheatgrass Seed Production 1. *Agronomy Journal*, 51(9), 515-518. <https://doi.org/10.2134/agronj1959.00021962005100090002x>
- Imran, M., Shakeel, A., Azhar, F., Farooq, J., Saleem, M., Saeed, A., Nazeer, W., Riaz, M., Naeem, M., & Javaid, A. (2012). Combining ability analysis for within-boll yield components in upland cotton (*Gossypium hirsutum* L.). *Genetics and Molecular Research* 11, 2790-2800.
- Jarwar, A. H., Wang, X., Iqbal, M. S., Sarfraz, Z., Wang, L., Ma, Q., & Shuli, F. J. P. J. B. (2019). Genetic divergence on the basis of principal component, correlation and cluster analysis of yield and quality traits in cotton cultivars. *Pak. J. Bot*, 51(3), 1143-1148.
- Li, Y., Si, Z., Wang, G., Shi, Z., Chen, J., Qi, G., ... & Zhang, T. (2023). Genomic insights into the genetic basis of cotton breeding in China. *Molecular Plant*, 16(4), 662-677.
- Malik, W., Ashraf, J., Iqbal, M. Z., Ali Khan, A., Qayyum, A., Ali Abid, M., Noor, E., Qadir Ahmad, M., & Hasan Abbasi, G. (2014). Molecular Markers and Cotton Genetic Improvement: Current Status and Future Prospects. *The Scientific World Journal*, 2014, 1-15. <https://doi.org/10.1155/2014/607091>
- May, O.L., D.T. Bowman and D.S. Calhoun. 1995. Genetic diversity of US Upland cotton cultivars released between 1980 and 1990. *Crop Sci.*, 35: 1570-1574.
- Meredith, W. R., & Bridge, R. R. (1972). Heterosis and Gene Action in Cotton, *Gossypium hirsutum* L. 1. *Crop Science*, 12(3), 304-310. <https://doi.org/10.2135/cropsci1972.0011183x001200030015x>
- Moulherat, C., Tengberg, M., Haquet, J. F., & Mille, B. (2002). First evidence of cotton at Neolithic Mehrgarh, Pakistan: Analysis of mineralized fibres from a copper bead. *Journal of Archaeological Science*, 29(12), 1393-1401. <https://doi.org/10.1006/jasc.2001.0779>
- Nadeem, a., Shakeel, a., Imtiaz, m., Khalid, m., & Amjad, i. (2022). Genetic variability studies for yield and within boll yield components in cotton (*Gossypium hirsutum* L.). *Biological and clinical sciences research journal*, 2022(1). <https://doi.org/10.54112/bcsrj.v2022i1.197>
- Nazir, A., J. Farooq, A. Mahmood, M. Shahid and M. Riaz. 2013. Estimation of genetic diversity for CLCuV, earliness and fiber quality traits using various statistical procedures in different crosses of *Gossypium hirsutum* L. *Vestnik. Orel. Gau.*, 43:2-9.
- Saeed F, Shabbir R.S., Farooq J., Riaz M., and Mahmood K., 2015, Genetic Diversity Analysis for Earliness, Fiber Quality and Cotton Leaf Curl Virus in *Gossypium hirsutum* L. *Accessions, Cotton Genomics and Genetics*, Vol.6, No.2 1-7 (doi: 10.5376/cgg.2015.06.0002)

- Saeed, F., Farooq, J., Mahmood, A., Riaz, M., Hussain, T., & Majid, A. (2013). Estimation of genetic variability for cotton leaf curl virus disease, fiber quality and some morphological traits using multivariate analysis in exotic gossypium (diploid and tetraploid) species. *Journal of Agrobiology*, 30(2), 107-115.
- Sarwar, G., Nazir, A., Rizwan, M., Shahzadi, E., & Mahmood, A. (2021). Assessment of genetic diversity of cotton cultivars by using correlation and principal component analysis for CLCuD tolerance, yield and quality traits. *J. Agric. Res*, 59(1), 1-5.
- Seyoum, M., Du, X. M., He, S. P., Jia, Y. H., Pan, Z., & Sun, J. L. (2018). Analysis of genetic diversity and population structure in upland cotton (*Gossypium hirsutum* L.) germplasm using simple sequence repeats. *Journal of Genetics*, 97(2), 513-522. <https://doi.org/10.1007/s12041-018-0943-7>
- Shakeel, A., Talib, I., Rashid, M., Saeed, A., Ziaf, K., & Saleem, M. F. (2015). Genetic diversity among upland cotton genotypes for quality and yield related traits. *Pakistan Journal of Agricultural Sciences* 52, 73-77.
- Sneath, P.H.A. and R.R. Sokal. 1973 Numerical taxonomy: the principles and practice of numerical classification. W. F. Free-Man and Co. San Francisco, 573 pp.
- Tyagi, P., Gore, M. A., Bowman, D. T., Campbell, B. T., Udall, J. A., & Kuraparthi, V. (2014). Genetic diversity and population structure in the US Upland cotton (*Gossypium hirsutum* L.). *Theoretical and Applied Genetics*, 127(2), 283-295. <https://doi.org/10.1007/s00122-013-2217-3>
- Wendel, j. F., brubaker, c. L., & percival, a. E. (1992). Genetic diversity in *Gossypium hirsutum* and the origin of upland cotton. *American journal of botany*, 79(11), 1291-1310. <https://doi.org/10.1002/j.1537-2197.1992.tb13734.x>
- Zafar, M. M., Mustafa, G., Shoukat, F., Idrees, A., Ali, A., Sharif, F., ... & Li, F. (2022). Heterologous expression of cry3Bb1 and cry3 genes for enhanced resistance against insect pests in cotton. *Scientific Reports*, 12(1), 10878.
- Zhi, X.-y., Han, Y.-c., Li, Y.-b., Wang, G.-p., Du, W.- l., Li, X.-x., Mao, S.-c., & Lu, F. (2016). Effects of plant density on cotton yield components and quality. *Journal of integrative agriculture* 15, 1469-1479.
- Zhu, G., Li, W., Wang, G., Li, L., Si, Q., Cai, C., & Guo, W. (2019). Genetic basis of fiber improvement and decreased stress tolerance in cultivated versus semi-domesticated upland cotton. *Frontiers in Plant Science*, 10, 1572.