ASSOCIATION OF SIMPLE SEQUENCE REPEAT MARKERS WITH MORPHOLOGY, DISEASE INDEX, LEAF PHOTOSYNTHETIC PIGMENTS AND SEED COTTON YIELD IN *Gossypium hirsutum* L. UNDER FIELD CONDITIONS

I. Sajjad[1], M. Saeed[1*] and G. Sarwar[2]

[1]Department of Botany, Faculty of Life Sciences, Government College University Faisalabad, 38000-Faisalabad, Pakistan.
[2]Ayub Agricultural Research Institute (AARI), Faisalabad, Pakistan.

[*]Corresponding Author's Email: saeed_pbg@gcuf.edu.pk

ABSTRACT

Cotton (*Gossypium hirsutum* L.) is a significant agricultural crop, primarily grown commercially for its natural fiber, while cottonseed is further processed to produce vegetable oil as a valuable by-product. This study aimed to conduct a genome-wide association study in cotton (*Gossypium hirsutum* L.) for morphology, cotton leaf curl virus disease index, leaf photosynthetic pigments content, and seed cotton yield under field conditions. One hundred cotton genotypes were planted under randomized complete block design in the field during 2021, 2022, and 2023 and data were collected for morphological traits, leaf photosynthetic pigment content, cotton leaf curl disease index (%), and yield attributes. Significant differences were observed among cotton genotypes for all studied traits. Plant height ranged from 80-190 cm, 64-172.67 cm, 110-245 cm, respectively, whereas, seed cotton yield per plant ranged from 9.12-262.83 g, 57.42-109.89 g, 18.78-201.6 g during 2021, 2022, 2023, respectively. Seed cotton yield per plant showed a significant positive correlation ($P \leq 0.01$) with the number of nodes, number of sympodial branches, number of bolls per plant, and boll weight. Cotton genotypes with stable and lower values of chlorophyll *a* and total chlorophyll had significantly higher seed cotton yield per plant. This suggested that under changing environmental conditions, cotton genotypes with optimal vegetative growth would produce sustainable seed cotton yield per plant. Cotton genotypes were genotyped using 62 polymorphic simple sequence repeat markers. Six SSR markers were found significantly associated ($P \leq 0.001$; R^2 Marker $\geq$ 10%) with seed cotton yield per plant. These markers were BNL-150 (06D), BNL-3398 (03A), BNL-3638 (08D), CIR-253 (05A), NAU-2336 (02D) and NAU-2579 (10D); whereas, 8 SSR markers BNL-3398 (03A), BNL-4015 (04A), BNL-2609 (05A), CIR-253 (05A), BNL-150 (06D), BNL-1513 (08D), BNL-3638 (08D), and BNL-3779 (09A) were associated with plant morphology, chlorophyll contents and yield traits simultaneously under 2 or 3 years phenotypic data. These identified simple sequence repeat markers would accelerate marker-assisted breeding efforts to develop elite cotton cultivars for sustainable cotton production under changing environments.

Keywords: cotton; genome-wide association study; photosynthetic pigments; simple sequence repeats (SSRs); sustainable cotton production

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0

https://doi.org/10.36899/JAPS.2026.6.0151 Published first online October 09, 2026

INTRODUCTION

Cotton (*Gossypium hirsutum* L.) belongs to the Malvaceae family. It has been under cultivation for about 8000 years and has become a widely cultivated species because it produces long-fiber cells that are highly suited for manufacturing processes in the textile industry (Huang *et al*., 2021). Cotton plays a central role in economic development of Pakistan and contributes 0.4 percent in GDP and 1.8 percent in value addition (Anonymous, 2026). Cotton has a lifespan of about 210 days and is cultivated as an annual crop, with a habit of perennial indeterminate fruiting and growth (Qian *et al*., 2023). During its growth period, it is exposed to a number of biotic and abiotic challenges in the field. The multitude of complex climatic adversities is the major concern challenging sustainable cotton production in Pakistan (Arshad *et al*., 2021; Kumar *et al*., 2023; Saeed *et al*., 2024). In cotton producing regions of Pakistan, some months experience water scarcity for crop production, followed by torrential rains. Both of these conditions are damaging for cotton production. Water-deficit damages the photosynthetic machinery and reduces chlorophyll *a* and *b* levels, quantum yield of photosystem-II, and dry matter accumulation (Wang *et al*., 2007; Zahid *et al*., 2021; Garrido *et al*., 2023). Under water-deficit, ethylene and other hormonal changes trigger flower and boll drop; the net effect is fewer,

lighter bolls and a large drop in seed-cotton yield. Drought-resilient genotypes show morphological and physiological traits that preserve boll set (Shani *et al*., 2025), leading to higher seed-cotton yield. Cotton is also exposed to high-temperature stress in the field as temperatures during the cotton growth period exceed 40°C (Zahid *et al*., 2016). This high-temperature stress leads to excessive shedding of squares, flowers, and young bolls (Zahid *et al*., 2016; Farooq *et al*., 2023). High daytime/nighttime temperatures (especially heat episodes during flowering) reduce pollen viability, disrupt anther dehiscence and increase flower/boll abscission, lowering boll number even when vegetative growth is maintained. Several recent studies report heat-induced flower sterility and greater boll shedding as primary causes of yield loss (Zahid *et al*., 2016; Bista *et al*., 2025; Ul-allah *et al*., 2026).

Plant morphology is important in light interception and whole-plant photosynthesis, supporting more fruiting sites and higher boll set. Plant height, number of sympodial branches, number of bolls per plant, boll weight, and seed cotton yield are good indicators of climate resilience (Saeed *et al*., 2024; Gören and Tan, 2024; Bista *et al*., 2025). Photosynthesis is an important attribute in plant development and stress tolerance. Chlorophyll *a* and *b* are essential photosynthetic pigments that enable plants to capture solar energy required for photosynthesis. Chlorophyll *a* is the primary photosynthetic pigment and participates directly in the reaction centers of Photosystems I and II, whereas chlorophyll *b* functions mainly as an accessory pigment that broadens the range of light absorbed and transfers excitation energy to chlorophyll *a* (Lokstein *et al*., 2021; Hunter and Morey-Burrows, 2026). The total chlorophyll content and Chl *a*/*b* ratio are, therefore, important indicators of light-harvesting capacity and photosynthetic performance, although photosynthesis is also influenced by factors such as light intensity, CO_2 availability, temperature, and water status (Simkin *et al*., 2022). Consequently, reductions in chlorophyll content under environmental stress can impair light absorption and energy conversion, potentially reducing photosynthetic efficiency and plant productivity (Stirbet *et al*., 2020). Cotton leaf curl virus disease (CLCuVD) is a major biotic constraint in Pakistan (Saeed *et al*., 2018). It is a viral disease transmitted by whitefly (*Bemisia tabaci*). Infected plants exhibit leaf curling, stunting, and reduced boll formation. The use of resistant varieties and control of whitefly populations are essential for managing CLCuVD (Afzal *et al*., 2023).

The weather uncertainty calls for an overall transformation of the cotton breeding program in Pakistan. Changed environmental conditions demand the development of climate-resilient cotton cultivars (Ul-allah *et al*., 2026). During the development of cotton cultivars, efforts should be made to pyramid genomic regions ensuring tolerance against multiple stresses simultaneously for sustainable cotton production in view of climate uncertainty. Profiling cultivars at genetic, biochemical, and morphological levels is necessary to characterize their potential for climate resilience and sustainable economic production (Shahnawaz *et al*., 2025). The molecular breeding approach, accelerated by genome-wide association studies (GWAS) and facilitated by marker-assisted selection (MAS), is very promising for developing cotton cultivars with desired traits in the shortest possible time (Sun *et al*., 2023; Ul-allah *et al*., 2026). GWAS, which links natural genetic variation (SSRs/SNPs/haplotypes etc.) to phenotypes in diverse germplasm, is a powerful tool to discover loci and candidate genes controlling stress tolerance and important agronomic traits (Sun *et al*., 2023; Kumar *et al*., 2024). At present, GWAS is widely used in plant sciences (Bilal *et al*., 2015; Kumar *et al*., 2024), especially in cotton (Kumari *et al*., 2022; Sun *et al*., 2023). Integrating GWAS discoveries with functional genomics, marker validation, genomic selection and marker-assisted breeding pipelines can substantially shorten the path from gene discovery to climate-resilient cultivar release. Marker-assisted selection (MAS) enables researchers to select traits of interest precisely through molecular analyses leading to superior breeding achievements. The quantitative trait locus (QTL) mapping (both linkage and association) approaches, combined with MAS, have simplified breeding programs by letting users precisely add stress tolerance genes into elite cotton cultivars (Shea, 2023). Several crucial genes, that control osmotic adjustment, boost antioxidant activity, and coordinate stress signaling pathways, have been successfully identified in the mechanism of drought tolerance (Luo *et al*., 2022). The MAS and QTL mapping are examples of genome technologies that have been crucial in identifying and incorporating stress-tolerant traits into cotton breeding frameworks (Kumar *et al*., 2024). Combining genomics with traditional breeding techniques offers encouraging opportunities to improve water-use efficiency, reduce reliance on over-irrigation, and ensure steady cotton production even during drought (Shah *et al*., 2024). The advancement of agricultural practices is possible with QTL mapping, GWAS and MAS which help scientists create climate-resilient cotton lines with better agronomic traits (Yunus *et al*., 2025). Researchers worldwide focus on developing climate-resilient cotton varieties through genomics-assisted breeding methods because the weather uncertainty in agricultural cotton cultivation continues to increase (Ndudzo *et al*., 2024). There is dire need to identify genomic regions associated with multiple morphological, biochemical and seed-cotton yield traits which will accelerate the development of climate-resilient cotton cultivars. The present study was designed to identify genomic regions in cotton (*Gossypium hirsutum* L.) associated with morphology, disease index, leaf photosynthetic pigments, and seed-cotton yield under field conditions.

MATERIALS AND METHODS

Plant material: One hundred genotypes of cotton (*Gossypium hirsutum* L.) were collected from Cotton Research Station (CRS), Ayub Agriculture Research Institute (AARI), Faisalabad, Pakistan out of more than 500 true-to-type world cotton germplasm. The plant material was grown three times in the field during the cotton growing season on May 5, 2021; May 2, 2022; and May 8, 2023 at the Research Field of Government College University Faisalabad, Faisalabad 38000, Pakistan. The year 2021 was characterized by hot (38.9–33.3°C from May to October) and relatively dry conditions (rainfall of approximately 161 mm in July, while the other months were comparatively drier). The year 2022 was characterized by generally greater rainfall (particularly in July) and sustained high temperatures (reaching 41.3°C in May and about 36°C during July–September), and 2023 by comparatively moderate temperatures (36.4°C in May, 38.3°C in June, and 35.6°C in July) with substantial monsoon rainfall (concentrated mainly in July–September) during the cotton season. The plant material was grown in the same experimental field during three years. The soil analysis of the experimental field during 2021–2023 indicated a loam-textured soil with a consistent saturation percentage of 41% throughout the three years. Soil pH remained slightly alkaline, ranging from 7.8 in 2021 to 8.0 in 2022 and 7.9 in 2023. Electrical conductivity (EC) varied only slightly from 2.1 to 2.3 dS m^{-1}, indicating relatively stable soil salinity across the study period. Available sodium remained high but relatively consistent (450–460 mg kg^{-1}), whereas available potassium showed minimal variation (100–101 mg kg^{-1}). Available calcium ranged from 62 to 65 mg kg^{-1}, while available phosphorus varied between 8.0 and 8.8 mg kg^{-1} during the three years. Soil organic matter remained constant at 0.70%, indicating a stable but relatively low organic-matter status of the experimental field throughout 2021–2023. The experiment used a randomized complete block design (RCBD) with three replications. Plant-to-plant and row-to-row distances were 30 cm and 75 cm, respectively. There was a 90-cm distance between replications. There were 10 plants grown for each entry, and data from eight plants were recorded (by discarding data from two bordering plants for each entry). There was absence of detectable cross-pollination despite the close spacing of plots that might be attributed to the predominantly self-pollinating nature of cotton and reduced insect-pollinator activity following pesticide application (Van Deynze *et al*., 2005). Plant material was managed using recommended agronomic practices, including irrigation, fertilizer application, weed control, pest management, and intercultural operations throughout during three cotton growing seasons. During the first year, 24 acre-inches of irrigation water flowed to the experimental plot. The same procedure was used to manage irrigation during the second and third years. The fertilizer application was optimized at 204 kg N, 85 kg P_2O_5, and 93 kg K_2O ha^{-1} sourced from Urea, DAP, and SOP across key growth stages. For weed control, recommended dose of pendimethalin was applied within 24 hours of sowing, and then after cotton germination glyphosate was sprayed for escaped grass and broadleaf weeds, and manual inter-cultivation was also done occasionally. For pest management, the recommended doses of insecticides diafenthiuron, flonicamid, and pyriproxyfen for sucking pests (whiteflies, jassids, thrips), and spinetoram, and emamectin benzoate for pink bollworms and armyworms were applied when pests threshold levels were reached. All husbandry practices were applied uniformly to maintain optimal crop growth and development.

Characterization of plant material for morphology, disease index, photosynthetic pigments and yield attributes: Plant material was evaluated for morphology, including plant height (PH), number of nodes per plant (NN), number of monopodial branches per plant (MB), number of sympodial branches (SB); cotton leaf curl virus disease index (DI, %); photosynthetic pigments such as chlorophyll *a* (Chl *a*, mg/g FW), chlorophyll *b* (Chl *b*, mg/g FW), ratio of chlorophyll *a* to *b* (Chl *a*/*b*), total chlorophyll (TC, mg/g FW); and yield attributes including number of bolls per plant (NB), boll weight (BW, g), and seed-cotton yield per plant (SCY, g).

Plant height (PH, cm) was measured using a measuring scale. NN, MB, and SB were counted. Disease index (DI, %) was recorded following the method of Akhtar *et al*. (2010). For Chl *a*, Chl *b*, TC, and Chl *a*/*b* measurement, fresh leaves weighing 0.5 grams were broken into small pieces and mixed with a 10 ml (80% acetone) solution. The mixture was kept overnight. Then the solution was centrifuged at 10,000 rpm, and the absorbance was determined by a spectrophotometer at 663, 645, and 480 nm. The chl *a*, chl *b*, Chl *a*/*b*, and TC were calculated using the following formulas:

Chl *a* (mg g^{-1} FW) = $[12.7(OD_{663}) - 2.69(OD_{645})] \times V/1000 \times W$

Chl *b* (mg g^{-1} FW) = $[22.9(OD645) - 4.68\ (OD663)] \times V/1000 \times W$

Where FW was the fresh weight of the leaf sample (g), V was the volume of the extract (ml), and OD was the optical density measured by the spectrophotometer (Arnon, 1949; Zhao *et al*., 2016). Chl *a*/*b* was calculated by dividing Chl *a* by Chl *b*, whereas TC was estimated by adding Chl *a* and Chl *b*. NB was measured by counting, whereas BW and SCY were measured by weighing with a physical balance.

Genotyping of the plant material with simple sequence repeats (SSR) markers: For genotyping of the cotton genotypes with SSR markers, 1-2 young leaves from each cotton genotype were collected from the field during 2021, transported to the lab by putting in ice box, and grounded in liquid nitrogen for DNA extraction. DNA was extracted following the CTAB method (Paterson *et al*., 1993). Quantification of the DNA was carried out using a NanoDrop™ 2000 Spectrophotometer, Thermo Scientific™, USA. DNA of the cotton genotypes was amplified using primers of SSR markers. DNA was amplified using primers in a PCR reaction solution with a final volume of 10 µl. DreamTaq Green PCR Master Mix (2X), Thermo Scientific™, USA was used to prepare the PCR reaction solution with the composition: 5 µl DreamTaq Green PCR Master Mix (2X); 0.5 µM forward primer; 0.5 µM reverse primer; 10 ng template DNA; nuclease-free water (to make a total volume of 10 µl). The program for the PCR reaction was set at 95 °C for 3 minutes; 30 cycles of (94 °C for 30 seconds, 50-57ºC for 30 seconds, 72 °C for 45 seconds); 72 °C for 5 minutes; and 4 °C for 10 minutes. PCR products were separated through polyacrylamide gel electrophoresis, stained with silver staining, and visualized by a gel documentation system. PCR product bands were recorded according to the sizes in base pairs (bp). A subset of 10 cotton genotypes were scored for polymorphism using 500 SSR markers. Out of these 500 SSR markers, 62 were found to be polymorphic and used to genotype the whole 100 cotton genotypes.

Structure and tassel analyses: The pattern of population/genetic structure in plant material was assessed using the software package STRUCTURE (Pritchard *et al*., 2000). It is a model-based clustering method. Genetic diversity, linkage disequilibrium (LD), and marker-trait associations were assessed by incorporating genotypic and phenotypic data; and Q (structure relatedness) matrices into TASSEL 2.1 software (Bradbury *et al*., 2007). The TASSEL stands for Trait Analysis for aSSociation, Evolution, and Linkage. The mean data of each trait was used separately for three years in TASSEL analysis.

Statistical analysis of data: The CoStat v. 6.303 was used (CoHort Software, 2004) to assess descriptive statistics, analysis of variance (ANOVA) estimates and the Pearson correlation coefficients for the studied traits. Two-way ANOVA was performed according to RCBD with three replications and two factors (100 genotypes, 3 years). The homogeneity of variance was tested using Levene's test (1960), and the data met the homogeneity assumption required for ANOVA. Principal component analysis (PCA), heatmap and trait correlation analyses were performed by using R-Studio software (Posit team, 2026). For these analyses combined mean of three years' data was used.

RESULTS

Descriptive statistics for the studied traits: Cotton genotypes displayed significant variation among all the studied traits during 3 years (Table 1). The PH ranged from 80-190 cm, 64-172.67 cm, 110-245 cm, whereas, SCY ranged from 9.12-262.83 g, 57.42-109.89 g, 18.78-201.6 g during 2021, 2022, 2023, respectively. Similarly, NB ranged from 4-74, 18-31, 6-60, whereas, BW ranged from 1.1-5.8 g, 3.06-3.82 g, and 1.78-4.05 g during 2021, 2022, 2023, respectively.

Table 1. Descriptive statistics for the studied traits.

Traits/Statistics	Mean	Sta. Dev.	Minimum	Maximum
PH (cm)	132.83	45.78	64	245
NN	5.62	1.90	1	13
MB	1.62	0.93	0.75	5
SB	24.42	5.86	5	44
DI (%)	12.03	12.23	0.32	56.55
NB	24.39	8.60	4	74
BW (g)	3.08	0.54	1.1	5.8
SCY (g)	74.41	27.02	9.12	262.83
Chl *a* (mg/g F. wt)	2.13	0.89	0.33	3.29
Chl *b* (mg/g F. wt)	0.52	0.25	0.04	2.58
Chl *a*/*b*	5.20	5.45	0.58	62.11
TC (mg/g F. wt)	2.66	0.97	0.62	5.00

PH plant height; *NN* number of nodes per plant; *MB* number of monopodial branches; *SB* number of sympodial branches; DI disease index (%); *NB* number of bolls per plant; *BW* boll weight; *SCY* seed-cotton yield per plant; *Chl a* chlorophyll *a*; *Chl b* chlorophyll *b*; *Chl a/b* ratio of chlorophyll *a* over chlorophyll *b*; *TC* total chlorophyll; *Sta. Dev* standard deviation.

ANOVA for the studied traits: Characterization of plant material for morphology, disease index, photosynthetic pigments, and yield attributes showed significant differences ($P \leq 0.001$) among genotypes for all the studied traits (Table 2). The values across years for the studied traits also differed significantly ($P \leq 0.001$), except for NB. The values for genotype × year interaction also differed significantly ($P \leq 0.001$) among genotypes for all the studied traits (Table 2).

Table 2. Mean squares of ANOVA for the studied traits.

Traits/SoV	PH	NN	MB	SB	DI (%)	NB	BW	SCY	Chl *a*	Chl *b*	Chl *a*/*b*	TC
Block	532.70 ns	2.06ns	1.11ns	1.91 ns	0.02ns	72.23*	0.01ns	503.73ns	0.03ns	0.02ns	2.14 ns	0.07 *
Genotype	1184.91 ***	4.49 ***	0.82 ***	36.97 ***	816.29 ***	184.94 ***	0.53 ***	1485.97 ***	0.30 ***	0.19 ***	103.32 ***	0.44 ***
Year	733751.77 ***	272.22 ***	93.14 ***	3866.80 ***	131.51 ***	13.75 ns	15.37 ***	13617.43 ***	325.01 ***	2.15 ***	1544.66 ***	369.81 ***
Genotype × Year	965.85 ***	4.71 ***	1.03 ***	36.09 ***	259.01 ***	171.67 ***	0.52 ***	1592.51 ***	0.11 ***	0.16 ***	61.45 ***	0.23 ***
Error	179.37	2.23	0.51	20.61	3.34	23.41	0.12	277.40	0.01	0.01	2.01	0.02
R^2	0.94	0.59	0.61	0.60	0.99	0.79	0.72	0.75	0.99	0.944	0.95	0.99
Coefficient of variation (%)	10.08	26.59	44.11	18.59	15.19	19.83	11.38	22.38	5.38	13.87	27.25	5.09

SoV Source of variation; *PH* plant height; *NN* number of nodes per plant; *MB* number of monopodial branches; *SB* number of sympodial branches; DI disease index (%); *NB* number of bolls per plant; *BW* boll weight; *SCY* seed-cotton yield per plant; *Chl a* chlorophyll *a*; *Chl b* chlorophyll *b*; *Chl a/b* ratio of chlorophyll *a* over chlorophyll *b*; *TC* total chlorophyll; R^2 SSmodel/SStotal.

Significant differences among blocks for NB and TC may reflect inherent spatial and micro-environmental heterogeneity across the experimental field. Differences in soil properties, moisture and nutrient availability, and microclimatic conditions may have contributed to this variation. The use of RCBD effectively accounted for such block-to-block variability, allowing more reliable comparison of cotton genotypes.

PCA biplot and Heatmap clustering of studied traits and cotton genotypes: The principal component analysis (PCA) biplot revealed distinct associations among the studied traits of cotton genotypes based on their contributions to the total variability (Figure 1). It illustrated the inter-relationships among various morphological, physiological, biochemical and yiled-related traits of cotton genotypes. In this biplot, the first two principal components (Dim1 and Dim2) explained 23.9% and 15.9% of the total variation, respectively, together summarizing the major sources of variability among the studied traits. Each arrow denoted a variable, with its direction and length representing the orientation and magnitude of its contribution to the principal components.

Traits such as TC, NB, and SCY showed strong positive correlations with the first principal component and exhibited high contribution values, as indicated by their red to orange color gradients. These variables were closely clustered, suggesting that yield-related traits were strongly interlinked with chlorophyll content, implying that genotypes with higher chlorophyll concentration tend to produce a greater number of bolls and higher yield potential. Similarly, Chl *a* and SB were moderately associated with Dim1, forming a secondary cluster near yield parameters. Their intermediate contribution (green shade) indicates a supportive role in influencing productivity through improved photosynthetic efficiency and plant architecture. Similarly, PH, also positively associated with Dim1, showed a lighter contribution, reflecting its moderate influence on yield traits under the given conditions. In contrast, traits such as BW, NN, MB, DI, and Chl *a/b* were positioned close to the origin, displaying weak correlations with both Dim1 and Dim2. Their blue to light green coloring pointed towards low contribution to the total variance, indicating that these parameters had a limited role in discriminating among the genotypes. On the other hand, Chl *b* showed a moderate correlation with TC and Chl *a*, contributing moderately to the overall variability (yellowish vector), suggesting its partial influence on the photosynthetic potential of the genotypes.

A) **B)**

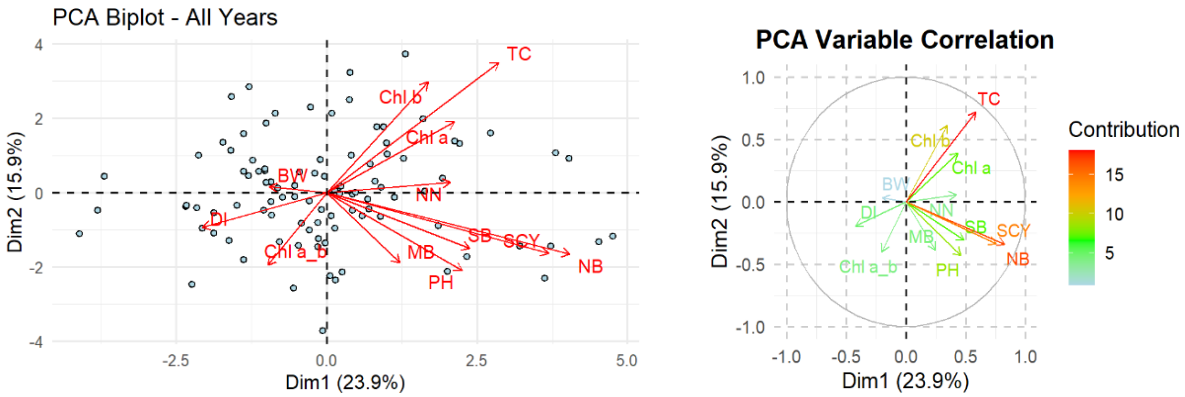

Figure 1. Principal component analysis based biplot showing studied traits (as vectors) and cotton genotypes (as cases): (A) PCA Biplot-All years; and (B) PCA Variable correlation. *PH* plant height; *NN* number of nodes per plant; *MB* number of monopodial branches; *SB* number of sympodial branches; DI disease index (%); *NB* number of bolls per plant; *BW* boll weight; *SCY* seed cotton yield per plant; *Chl a* chlorophyll *a*; *Chl b* chlorophyll *b*; *Chl a/b* ratio of chlorophyll *a* over chlorophyll *b*; *TC* total chlorophyll.

Overall, PCA separated the variables into three functional groups: (i) yield and chlorophyll-related traits (TC, SCY, NB) with high contribution; (ii) growth and photosynthetic parameters (Chl *a*, Chl *b*, SB, PH) with moderate contribution; and (iii) morpho-physiological and yield-related traits (NN, MB, DI, Chl *a/b*, BW) with minimal impact. This grouping pattern highlights that yield performance in cotton genotypes is primarily driven by chlorophyll content and number of bolls traits, emphasizing the importance of photosynthetic and yield-related parameters under the evaluated conditions.

The cumulative heatmap analysed standardized values of 12 morpho-physiological and yield-related traits (PH, NN, MB, SB, DI, NB, BW, SCY, Chl *a*, Chl *b*, Chl *a/b*, TC,) across the evaluated genotypes, with hierarchical clustering applied to both traits (rows) and genotypes (columns) (Figure 2). The color scale ranged from cool (green — lower standardized values) through neutral (yellow — average) to warm (orange — higher standardized values). Overall, the heatmap showed that most trait–genotype combinations fall within the mid-range (predominantly yellow), indicating that for many genotypes the traits were close to the experiment mean; super-high or super-low values were present but scattered rather than concentrated.

The row clustering (trait dendrogram) grouped traits with similar patterns across genotypes. From the visual pattern, chlorophyll-related traits (Chl *a*, TC, Chl *b*) and NN tended to cluster together, reflecting their joint physiological basis and similar variation among genotypes. This grouping supported earlier PCA results where chlorophyll measures and NN loaded together and contributed similarly to genotype separation. Yield-related traits such as (NB, SCY) and morphological traits (PH, MB, SB) also showed similar profiles across many genotypes and formed a cluster or subcluster, reinforcing the positive association between yield-related and morphological traits. Morphological traits (PH, SB, MB, NN) and disease index (DI) showed more heterogeneous patterns: some genotypes presented elevated plant morphological values while others do not, producing a more mixed color mosaic for those rows. Boll weight (BW) also displayed a relatively independent pattern, often not matching perfectly with either photosynthetic pigment or simple yield clusters, which suggested BW may be governed by different genetic or environmental factors than NB or SCY. The genotype (column) dendrogram, genotypes were separated into distinct clusters with small blocks of warmer color (orange) indicating genotypes that were relatively high for several traits simultaneously — these are candidate high-performers, particularly where elevated NB/SCY and higher TC/Chl coincided. Conversely, small green patches indicated genotypes with below-average standardized values for particular traits. Importantly, the heatmap suggested no single genotype is uniformly extreme across all traits; instead, genotypes showed trait-specific strengths and weaknesses, which implied breeding gains would require multi-trait selection rather than targeting a single trait.

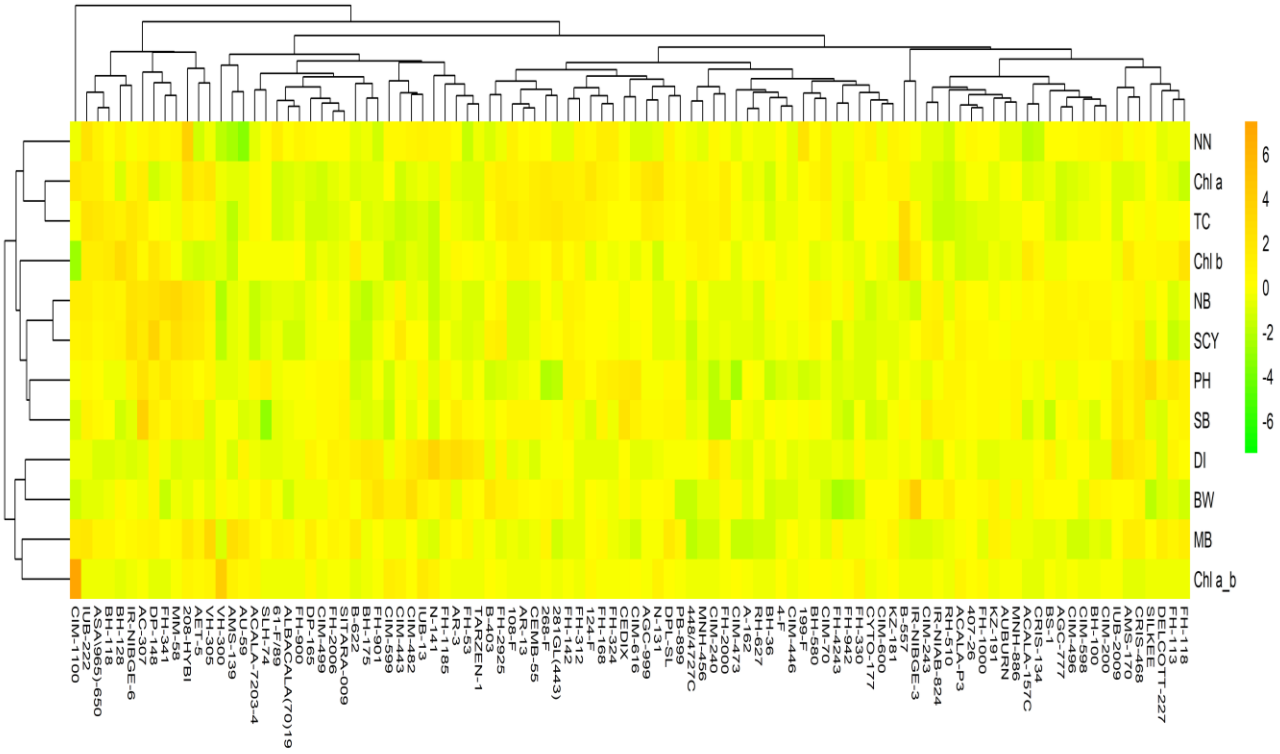

Figure 2. Heatmap of studied traits. *PH* plant height; *NN* number of nodes per plant; *MB* number of monopodial branches; *SB* number of sympodial branches; DI disease index (%); *NB* number of bolls per plant; *BW* boll weight; *SCY* seed cotton yield per plant; *Chl a* chlorophyll *a*; *Chl b* chlorophyll *b*; *Chl a/b* ratio of chlorophyll *a* over chlorophyll *b*; *TC* total chlorophyll.

These patterns confirmed considerable genetic variability and suggested that selection based on combined morpho-physiological indices could enhance yield stability and adaptability across environments. Scattered warm spots within chlorophyll and yield rows highlighted genotypes with particularly high photosynthetic pigment content and seed-cotton yield. These genotypes are of interest for selection. The prevalence of neutral (yellow) cells meant that much of the variation was moderate; therefore, combining this heatmap information with PCA loadings and cluster membership would better identify genotypes that balance high chlorophyll and yield with acceptable morpho-physiology and architecture. Finally, the dendrograms provided a visual basis for grouping genotypes for further comparative analysis (e.g., testing mean differences between genotype clusters, or computing selection indices that weight TC, NB and SCY more heavily).

Trait correlation: Differences were observed in the pattern of trait correlation and PCA and heatmap results. A moderate positive correlation was exhibited by PH with NN and a strong positive correlation with MB (Figure 3). It showed a negative association with SB and a negligible relationship with DI. The correlations of PH with NB, BW and SCY were weak and non-significant. Regarding biochemical traits, PH exhibited a strong positive association with Chl *a* and TC whereas its relationship with Chl *b* and Chl *a/b* remained weak. A moderate positive correlation was exhibited by NN with PH. Its correlation with MB, SB, NB, BW, and SCY were weak and non-significant, whereas, it showed weak or negative relationship with DI, Chl *b*, Chl *a/b*, and TC while its correlation with Chl *a* was also very weak. A strong positive correlation was shown by MB with PH, whereas, its correlation with NN was weak, while a negative relationship was observed with SB and DI. Weak associations were exhibited by MB with NB, BW, and SCY indicating limited contribution to yield components. A moderate to strong positive correlation was shown by MB with Chl *a* and TC while its relationship with Chl *b* and Chl *a/b* remained weak. Negative associations were exhibited by SB with PH and MB. Its correlation with NN and DI were very weak, whereas, SB showed weak positive correlation with NB, BW and moderate positive correlation with SCY. It showed

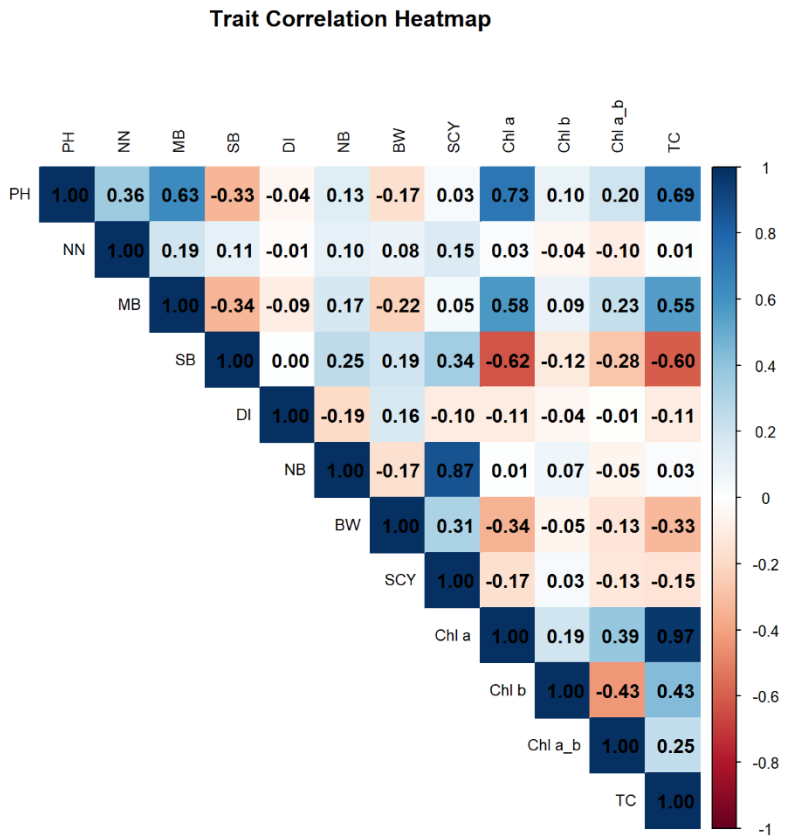

Figure 3. Correlation coefficients for the studied traits. *PH* plant height; *NN* number of nodes per plant; *MB* number of monopodial branches; *SB* number of sympodial branches; DI disease index (%); *NB* number of bolls per plant; *BW* boll weight; *SCY* seed cotton yield per plant; *Chl a* chlorophyll *a*; *Chl b* chlorophyll *b*; *Chl a/b* ratio of chlorophyll *a* over chlorophyll *b*; *TC* total chlorophyll.

strong negative correlation with Chl *a* and TC while its relationship with Chl *b* and Chl *a/b* remained weak to moderately negative. Weak negative associations were exhibited by DI with NB, SCY, Chl *a* and PH while its correlation with BW was weak positive. Its association with NN, MB, SB, Chl *b*, Chl *a/b* ratio and TC was non-significant negligible. A very strong positive correlation was exhibited by NB with SCY. It showed weak positive association with PH, NN, MB and SB. It had weak negative correlation with DI and BW while its relationship with biochemical traits Chl *a*, Chl *b*, Chl *a/b* and TC was weak. A very strong positive correlation was exhibited by BW with SCY indicating its major contribution to overall yield. It showed weak negative association with PH and MB and a weak positive correlation with SB and DI. It had very weak correlation with NN, NB and biochemical traits Chl *a*, Chl *b*, Chl *a/b* ratio and TC. A very strong positive correlation was exhibited by SCY with NB. It also showed moderate positive association with SB and BW. It showed weak positive correlation with NN, PH and MB while its relationship with DI was weak negative. It showed a weak to negligible negative correlation with Chl *a*, Chl *b*, Chl *a/b* and TC. Chlorophyll *a* showed a very strong positive correlation with TC indicating its major contribution to overall pigment content. It also showed strong positive association with PH and MB. It had a strong negative correlation with SB and a weak negative relationship with BW and DI. Its correlation with NN, NB and SCY was negligible, while a weak to moderate positive association was observed with Chl *b* and Chl *a/b*. Chlorophyll *b* exhibited a moderate positive correlation with TC. It showed weak positive association with PH and MB and weak negative correlation with SB, BW and DI. It had negligible association with NN, NB and SCY. A weak positive association was observed with Chl *a* while a moderate negative correlation occurred with Chl *a/b*. Chlorophyll *a/b* exhibited a moderate positive correlation with Chl *a* and TC. It showed a moderate negative correlation with Chl *b*. It displayed weak positive association with PH and MB while correlation with SB was weak negative. Its association with NN, DI, NB, BW and SCY was negligible. A very strong positive correlation was exhibited by TC with Chl *a* and moderate positive correlation with Chl *b* indicating that both pigments contribute to overall chlorophyll content, with Chl *a* being the major component. The association of TC was strong positive with PH and MB, whereas, the correlation of TC with NN, DI, NB, BW and SCY was weak or negligible.

Identification of significant marker-trait associations: Significant marker-trait associations ($P \leq 0.001$; R^2 Marker $\geq$ 10%) were identified in this study (Table 3). Three SSR markers were found to be associated with PH. These markers were BNL-2715, HAU-3434, and NAU-2579. These markers were located on chromosomes 05D, 10D, and 10D, respectively.

Table 3. Marker-trait associations identified for the studied traits ($P \leq 0.001$; R^2 Marker $\geq$ 10%).

Trait	Locus	Chr	R^2-Marker (%)
NN, MB, SB, BW, SCY, Chl *a*, Chl *b*	BNL-3638	08D	28.06 - 82.65
NN, MB, SB, NB, Chl *a*, Chl *b*, TC	BNL-2609	05A	28.69 - 63.21
MB, NB, BW, SCY, Chl *a*, Chl *b*	BNL-3398	03A	16.26 - 77.36
MB, NB, BW, SCY, Chl *a*, Chl *b*	BNL-150	06D	17.28 - 57.61
NN, MB, BW, Chl *a*, Chl *b*	BNL-4015	04A	25.54 - 87.62
MB, BW, Chl *a*, Chl *b*	BNL-3779	09A	55.54 - 87.35
NN, MB, Chl *a*, Chl *b*	BNL226	03A	35.55 - 97.61
PH, MB, Chl *a*, Chl *b*	BNL-2715	05D	47.36 - 77.25
MB, SCY, Chl *a*, Chl *b*	CIR-253	05A	53.2 - 91.52
NN, BW, Chl *a*	BNL-1513	08D	35.24 - 89.63
NB, SCY, TC	NAU2336	02D	19.52 - 59.84
PH, SB, NB, SCY	NAU-2579	10D	27.6 - 70.74
SB, TC	TMB-0409	07A	27.53 - 75.47
PH, TC	HAU-3434	10D	22.21 - 70.71
BW, Chl *b*	HAU-0438	02D	69.76 - 87.52
NN, BW	NAU1103	11A	69.78 - 89.64

PH plant height; *NN* number of nodes per plant; *MB* number of monopodial branches; *SB* number of sympodial branches; *NB* number of bolls per plant; *BW* boll weight; *SCY* seed-cotton yield per plant; *Chl a* chlorophyll *a*; *Chl b* chlorophyll *b*; *TC* total chlorophyll; R^2 phenotypic variance explained.

Nine markers were found to be associated with MB. These markers were BNL-150 (06D), BNL-226 (03A), BNL-2609 (05A), BNL-2715 (05D), BNL-3398 (03A), BNL-3638 (08D), BNL-3779 (09A), BNL-4015 (04A) and CIR-253 (05A). Four markers were found associated with SB. These markers were BNL-2609 (05A), BNL-3638 (08D), NAU-2579 (10D) and TMBO-0409 (07A). Five markers were found to be associated with NB. These markers were BNL-150 (06D), BNL-2609 (05A), BNL-3398 (03A), NAU-2336 (02D) and NAU-2579 (10D). Eight markers were found associated with BW. These markers were BNL-150 (06D), BNL-1513 (08D), BNL-3398 (03A), BNL-3779 (09A), BNL-4015 (04A), HAU-0438 (02D), BNL-150 (06D) and NAU-1103 (11A). Six SSR markers were found to be significantly associated ($P \leq 0.001$; R^2-Marker $\geq$ 10%) with SCY. These markers were BNL-150 (06D), BNL-3398 (03A), BNL-3638 (08D), CIR-253 (05A), NAU-2336 (02D) and NAU-2579 (10D). Ten markers were found to be associated with Chl *a*. These markers were BNL-150 (06D), BNL-1513 (08D), BNL-226 (03A), BNL-2609 (05A), BNL-2715 (05D), BNL-3398 (03A), BNL-3638 (08D), BNL-3779 (09A), BNL-4015 (04A) and CIR-253 (05A). Ten markers were found to be associated with Chl *b*. These markers were BNL-150 (06D), BNL-226 (03A), BNL-2609 (05A), BNL-2715 (05D), BNL-3398 (03A), BNL-3638 (08D), BNL-3779 (09A), HAU-0438 (02D), BNL-4015 (04A) and CIR-253 (05A). Three markers were found to be associated with TC. These markers were HAU-3434 (10D), NAU-2336 (02D) and TMBO-0409 (07A).

From the above findings, it was evident that chromosomes 03A, 04A, 05A, 06D, 08D, and 09A contained important genomic regions involved in photosynthesis and better vegetative growth of cotton plants, resulting in better SCY. The SSR markers BNL-3398 (03A), BNL-4015 (04A), BNL-2609 (05A), CIR-253 (05A), BNL-150 (06D), BNL-1513 (08D), BNL-3638 (08D), and BNL-3779 (09A) which were associated with plant morphology, chlorophyll contents and yield traits simultaneously under 2 or 3 years' phenotypic data would be promising candidates for a marker-assisted breeding (MAB) approach to develop elite cotton cultivars for sustainable cotton production.

DISCUSSION

Climate change (drought, heat, salinity, erratic rainfall) increasingly threatens global cotton production (Saeed *et al*., 2024). Climate change is characterized by unpredictable weather patterns. To ensure sustainable production amidst this uncertainty, crop cultivars must possess multiple sources of climate resilience. The present research utilized various

biochemical, molecular, morphological, and yield-related attributes to assess the potential of cotton for improved seed-cotton yield under field conditions. Plant material was characterized using PH, NN, MB, SB, DI (%), BP, BW, SCY, Chl *a*, Chl *b*, Chl *a*/*b*, and TC, and SSR markers linked to these traits were identified through GWAS.

SSR markers have been extensively applied in GWAS to dissect the genetic architecture of important cotton plant morphology traits, including plant height, number of main-stem nodes, sympodial branches, and monopodial branches. These traits are biologically interconnected because cotton plant height depends largely on node production and internode elongation, whereas the number and distribution of monopodial and sympodial branches determine canopy structure, reproductive capacity, planting-density adaptation, and ultimately yield (Chang *et al*., 2025). In the present study, 14 SSR markers were found associated with plant morphology (PH, NN, SB, and MB) across 2 or 3 years. These markers were BNL-226 (03A), BNL-3398 (03A), BNL-4015 (04A), BNL-2609 (05A), CIR-253 (05A), BNL-2715 (05D), BNL-150 (06D), TMB-0409 (07A), BNL-1513 (08D), BNL-3638 (08D), BNL-3779 (09A), NAU-2579 (10D), HAU-3434 (10D), and NAU1103 (11A). The phenotypic variance explained (R^2) value for these markers ranged from 16.26% - 97.61%. The detection of stable associations across years demonstrated the potential of SSR-based GWAS for identifying reproducible loci affecting plant morphology. In a representative SSR-based association-mapping study, 172 upland cotton accessions were phenotyped for five plant-architecture traits across multiple environments and genotyped with 386 SSR markers, resulting in 66 significant marker–trait associations, of which 35 were repeatedly detected in more than one environment (Li *et al*., 2016). Association mapping using SSR markers has also identified 20 highly significant quantitative trait SSRs for fruit-branch node number and branch length across three environments, confirming that branch architecture is influenced by additive, dominance, epistatic, and environment-interaction effects (Mei *et al*., 2017). A recent GWAS report revealed the dynamic genetic basis of plant height in upland cotton and demonstrated that different genomic loci can contribute to plant height at different developmental stages (Ye *et al*., 2023). Collectively, comparing and analysing the results of present study and the previous reports together indicated that plant height, node number, and sympodial branch development are highly polygenic and developmentally regulated traits (Chang *et al*., 2025). SSR loci associated with general plant morphology and branch traits provide useful starting points for identifying genomic regions influencing the balance between vegetative monopodial growth and reproductive sympodial branching. Overall, SSR-based association mapping, together with recent SNP-based GWAS, provides a complementary molecular framework for improving plant height, node number, and branch architecture and for developing cotton cultivars with compact, productive, and mechanization-friendly plant types.

SSR-based association mapping has been used to identify allelic variation associated with leaf chlorophyll traits and to support molecular breeding for improved photosynthetic performance. Liu *et al*. (2017) analyzed natural population structure and marker-trait associations between leaf chlorophyll content and 137 SSR markers. This analysis detected 355 polymorphic alleles, with an average of 2.6 alleles per SSR locus, demonstrating sufficient genome-wide diversity for association analysis. Using a general linear model (GLM), 22 alleles showed significant associations with leaf chlorophyll content and individually explained 5.28–10.85% of the phenotypic variation. The alleles SWU0529a and NAU998c showed the largest phenotypic effects in the GLM analysis, explaining 10.85% and 10.48% of the variation, respectively. A more stringent mixed linear model (MLM) identified 17 significant alleles that explained 3.72–8.58% of phenotypic variation, with SWU0923b and SWU0662d among the strongest associations. Importantly, 12 alleles were significant in both GLM and MLM analyses, providing greater confidence that these SSR associations were not solely caused by population structure or relatedness. NAU998c was consistently associated with chlorophyll content at three sampling stages in both analytical models, indicating a particularly stable marker-trait relationship. This study primarily evaluated overall leaf chlorophyll rather than separately reporting marker associations for chlorophyll *a*, chlorophyll *b*, and total chlorophyll; therefore, direct SSR-GWAS evidence resolving each pigment fraction independently remained comparatively limited. In the present study, 14 SSR markers were found associated with chlorophyll content (Chl *a*, Chl *b*, or TC). These markers were NAU2336 (02D), HAU-0438 (02D), BNL-226 (03A), BNL-3398 (03A), BNL-4015 (04A), BNL-2609 (05A), CIR-253 (05A), BNL-2715 (05D), BNL-150 (06D), TMB-0409 (07A), BNL-1513 (08D), BNL-3638 (08D), BNL-3779 (09A), and HAU-3434 (10D). Phenotypic variance explained (R^2) value for these markers ranged from 16.26% - 97.61%. A 2026 review of molecular-marker and GWAS approaches in cotton emphasized the continuing value of integrating historical marker associations with modern genome-wide and functional-genomic resources for trait improvement (Chen *et al*., 2026). Consequently, SSR markers identified in this study represent useful candidates for MAS to accelerate development of climate-resilient cotton cultivars.

The GWAS has been effectively used to identify marker-trait associations for major cotton yield components, including number of bolls per plant, boll weight, and seed cotton yield. A recent and comprehensive SSR-based study by Kumar *et al*. (2021) evaluated upland cotton under early- and normal-sown conditions during two seasons and identified numerous significant marker-trait associations for lint-yield contributing characters, including number of bolls, boll weight, and seed-cotton yield per plant. Among the promising associations detected in at least two environments, CGR5732 was significantly associated with number of bolls per plant and seed-cotton yield per plant, whereas BNL3257

was also associated with both boll number and seed-cotton yield. The SSR marker BNL3279 showed association with number of bolls per plant, seed-cotton yield per plant, and lint yield, indicating either pleiotropic effects or close linkage among loci governing related yield traits. For boll weight, BNL3976 was identified as a significant marker-trait association and was also associated with seed index, suggesting a genomic relationship between boll development and seed-related components. Other SSR markers, including BNL3452, BNL3441, BNL3479, and BNL1721, showed associations with number of bolls, seed-cotton yield, or multiple yield-related traits depending on the sowing environment, emphasizing the complex genetic architecture of cotton productivity. In the present study, 12 SSR markers were found associated with NB, BW or SCY in 2 or 3 years' phenotypic data. These markers were NAU2336 (02D), HAU-0438 (02D), BNL-3398 (03A), BNL-4015 (04A), BNL-2609 (05A), CIR-253 (05A), BNL-150 (06D), BNL-1513 (08D), BNL-3638 (08D), BNL-3779 (09A), NAU-2579 (10D) and NAU-1103 (11A). Phenotypic variance explained (R^2) values for these markers were high and ranged from 16.26% - 91.52%. These findings are consistent with modern high-density GWAS studies, which continue to reveal large numbers of quantitative trait loci and candidate genes for boll number, boll weight, and yield traits confirming that these traits are highly polygenic (Zhu *et al*., 2021; Shi *et al*., 2024; Shui *et al*., 2024). These studies provided independent genomic confirmation that the yield traits are governed by multiple loci, including small-effect and environmentally responsive loci. The repeated detection of SSR associations across years in our study is particularly valuable because stable marker-trait associations are more promising for MAS than associations detected in a single experiment. Specifically, 8 SSR markers, BNL-3398 (03A), BNL-4015 (04A), BNL-2609 (05A), CIR-253 (05A), BNL-150 (06D), BNL-1513 (08D), BNL-3638 (08D), and BNL-3779 (09A) which were associated with plant morphology, chlorophyll contents and yield traits simultaneously under 2 or 3 years' phenotypic data are promising candidates for MAS to develop elite climate-resilient cotton cultivars.

Conclusion: Based on the findings of this study, it was concluded that it was possible to pyramid genomic loci associated with desirable plant morphology, chlorophyll contents, and seed cotton yield traits in a single cotton cultivar as a number of promising significant SSR markers were identified associated with multiple traits simultaneously in 2 or 3 years' phenotypic data. The phenotypic variance explained (R^2) values for these associations were high enough to reveal their potential as promising candidates for future MAS programs to develop elite cotton cultivars with sustainable cotton production under post climate change scenario. These promising SSR markers were BNL-3398 (03A), BNL-4015 (04A), BNL-2609 (05A), CIR-253 (05A), BNL-150 (06D), BNL-1513 (08D), BNL-3638 (08D), and BNL-3779 (09A). These can be employed in functional genomics studies to elucidate their detailed role and mechanism in seed cotton yield enhancement.

Author Contributions IS performed the research work and wrote manuscript; MS conceived the project, designed the experiments, analysed the results, and edited manuscript; and GS edited manuscript.

Conflict of Interest Statement All authors declare that there is no conflict of interest.

Acknowledgements The authors would like to acknowledge the kind support of the Cotton Research Station (CRS), Ayub Agricultural Research Institute (AARI), Faisalabad, Pakistan, for providing guidance on agronomic practices during this study.

REFERENCES

Afzal, M., S. Saeed, H. Riaz, M. Ishtiaq and M.H. Rahman (2023). A critical review of whitefly (*Bemisia Tabaci* Gennadius) cryptic species associated with the cotton leaf curl disease. J. Innov. Sci. 9. https://doi.org/10.17582/journal.jis/2023/9.1.24.43

Akhtar K.P., S. Haidar, M.K.R. Khan, M. Ahmad, N. Sarwar, M.A. Murtaza and M. Aslam (2010). Evaluation of Gossypium species for resistance to cotton leaf curl Burewala virus. Ann. Appl. Biol. 157: 135-147. https://doi.org/10.1111/j.1744-7348.2010.00416.x

Anonymous (2026). Pakistan Economic Survey 2025-2026. Government of Pakistan, Finance Division, Islamabad, Pakistan, pp. 23. chrome- extension ://efaidnbmnnnibpcajpcglclefindmkaj/ https://www.finance.gov.pk/survey/chapter_26/Complete_PES2025_26.pdf

Arnon, D.I. (1949). Copper enzymes in isolated chloroplasts. Polyphenoloxidase in *Beta vulgaris*. Plant. Physiol. 24: 1-15. https://doi.org/10.1104/pp.24.1.1

Arshad A., M.A. Raza, Y. Zhang, L. Zhang, X. Wang, M. Ahmed and M. Habib-ur-Rehman (2021). Impact of climate warming on cotton growth and yields in China and Pakistan: A regional perspective. Agriculture. 11(2): 97. https://doi.org/10.3390/agriculture11020097

Bilal, M., M. Saeed, I.A. Nasir, B. Tabassum, M. Zameer, A. Khan, M. Tariq, M.A. Javed and T. Husnain (2015). Association mapping of cane weight and tillers per plant in sugarcane. Biotechnol. Biotechnol. Equip. 29(4): 617-623. https://doi.org/10.1080/13102818.2015.1008203

Bista, M.K., N. Kodadinne Narayana, A. Chakravaram, B. Pieralisi, J. Dhillon, K.R. Reddy and R. Bheemanahalli (2025). Intensifying heat stress impacts cotton flowering and boll development efficiency. BMC. Plant. Biol. 25: 984. https://doi.org/10.1186/s12870-025-06934-8

Bradbury, P.J., Z. Zhang, D.E. Kroon, T.M. Casstevens, Y. Ramdoss and E.S. Buckler (2007). TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics. 23(19): 2633-263. https://doi.org/10.1093/bioinformatics/btm308

Chang, H., H. Ji, R. Liu, J. Feng, J. Wang, S. Zhao, W. Li, Z. Qiu, N.I. Elsheery, S. Yu, L. Li and Z. Feng (2025). Genome-wide association study reveals significant loci and candidate genes for fruit branch length in upland cotton. Plant. Genome. 18(2):e70041. https://doi.org/10.1002/tpg2.70041

Chen, L., S. Li, X. Wang, L. Ke, Y. Yu and Y. Sun (2026). Advancing cotton improvement: a review of functional gene mining through molecular markers, GWAS, and map-based cloning. J. Cotton. Res. 9: 28. https://doi.org/10.1186/s42397-026-00272-2

CoHort Software. (2004). CoStat: Statistical analysis package, Version 6.303. CoHort Software, Monterey, CA, USA.

Farooq, A., A. Shakeel, A. Saeed, J. Farooq, M. Rizwan, W.S. Chattha, G. Sarwar and R. Yasir (2023). Genetic variability predicting breeding potential of upland cotton (*Gossypium hirsutum* L.) for high temperature tolerance. J. Cotton. Res. 6: 7. https://doi.org/10.1186/s42397-023-00144-z

Garrido, A., A. Conde, J. Serôdio, R.C.H. De-Vos and A. Cunha (2023). Fruit photosynthesis: more to know about where, how and why. Plants. 12(13): 2393. https://doi.org/10.3390/plants12132393

Gören, H.K. and U. Tan (2024). Unraveling the complexities of drought stress in cotton: a multifaceted analysis of selection criteria and breeding approaches. PeerJ. 12: e17584. https://doi.org/10.7717/peerj.17584

Huang, G., J.Q. Huang, X.Y. Chen and Y.X. Zhu (2021). Recent advances and future perspectives in cotton research. Annu. Rev. Plant. Biol. 72: 437-462. https://doi.org/10.1146/annurev-arplant-080720-113241

Hunter, C.N. and F.S. Morey-Burrows (2026). The structural chemistry and biosynthesis of chlorophylls. RSC. Chem. Biol. 7(7): 1134-1166. https://doi.org/10.1039/d6cb00082g

Kumar, P., S. Nimbal, R.S. Sangwan, N. Budhlakoti, V. Singh, D.C. Mishra, Sagar and R.R. Choudhary (2021). Identification of novel marker–trait associations for lint yield contributing traits in upland cotton (*Gossypium hirsutum* L.) using SSRs. Front. Plant. Sci. 12: 653270. https://doi.org/10.3389/fpls.2021.653270

Kumar, R., S.P. Das, B.U. Choudhury, A. Kumar, N.R. Prakash, R. Verma, M. Chakraborti, A.G. Devi, B. Bhattacharjee, R. Das, B. Das, H.L. Devi, B. Das, S. Rawat and V.K. Mishra (2024). Advances in genomic tools for plant breeding: harnessing DNA molecular markers, genomic selection, and genome editing. Biol. Res. 57: 80. https://doi.org/10.1186/s40659-024-00562-6

Kumar, R., S.K. Mishra, K. Singh, I. Al-Ashkar, M.A. Iqbal, M.N. Muzamil, M. Habib ur Rahman and A. El Sabagh (2023). Impact analysis of moisture stress on growth and yield of cotton using DSSAT-CROPGRO-cotton model under semi-arid climate. Peer. J. 11: e16329. https://doi.org/10.7717/peerj.16329

Kumari, M., L. Muduli, P.K. Meher and S.K. Pradhan (2022). Genome-wide association study (GWAS) for trait analysis in crops. In: Wani, S.H. and A. Kumar (eds) Genomics of Cereal Crops. Springer Protocols Handbooks. Humana, New York, NY.

Levene, H. (1960). Robust tests for equality of variances. In: Olkin, I. (Ed.), Contributions to probability and statistics: Essays in Honor of Harold Hotelling. Stanford University Press, Stanford, CA (USA), pp. 278–292.

Li, C.Q., N.J. Ai, Y.J. Zhu, Y.Q. Wang, X.D. Chen, F. Li, Q.Y. Hu and Q.L. Wang (2016). Association mapping and favourable allele exploration for plant architecture traits in upland cotton (*Gossypium hirsutum* L.) accessions. J. Agric. Sci. 154(4): 567-583. https://doi.org/10.1017/S0021859615000428

Liu, Q., L. Li, C. Zhang, J. Su, H. Wei, H. Wang and S. Yu (2017). Association analysis of leaf chlorophyll content with SSR markers and exploration of superior alleles in upland cotton. Sci. Agric. Sin. 50(18): 3439–3449. https://doi.org/10.3864/j.issn.0578-1752.2017.18.001

Lokstein, H., G. Renger, J.P. Götze (2021). Photosynthetic light-harvesting (Antenna) complexes-Structures and Functions. Molecules. 26(11): 3378. https://doi.org/10.3390/molecules26113378

Luo, C., W. Min, M. Akhtar, X. Lu, X. Bai, Y. Zhang and P. Li (2022). Melatonin enhances drought tolerance in rice seedlings by modulating antioxidant systems, osmoregulation, and corresponding gene expression. Int. J. Mol. Sci. 23(20): 12075. https://doi.org/10.3390/ijms232012075

Mei, Y., J. Yu, A. Xue, S. Fan, M. Song, C. Pang, W. Pei, S. Yu and J. Zhu (2017). Dissecting genetic network of fruit branch traits in upland cotton by association mapping using SSR markers. PLoS. One. 12(1): e0162815. https://doi.org/10.1371/journal.pone.0162815

Ndudzo, A., A.S. Makuvise, S. Moyo and E.D. Bobo (2024). CRISPR-Cas9 genome editing in crop breeding for climate change resilience: Implications for smallholder farmers in Africa. J. Agric. Food. Res. 16: 101132. https://doi.org/10.1016/j.jafr.2024.101132

Paterson, A.H., C.L. Brubaker and J.F. Wendel (1993). A rapid method for extraction of cotton (Gossypium spp.) genomic DNA suitable for RFLP or PCR analysis. Plant. Mol. Biol. Rep. 11: 122-127. https://doi.org/10.1007/BF02670470

Posit team. (2026). RStudio: Integrated development environment for R. Posit Software, PBC, Boston, MA, USA.

Pritchard, J.K., M. Stephens and P. Donnelly (2000). Inference of population structure using multilocus genotype data. Genetics. 155(2): 945-959. https://doi.org/10.1093/genetics/155.2.945

Qian, L., H.Y. Meng, X.H. Chen and R. Tang (2023). Evaluating agricultural drought and flood abrupt alternation: A case study of cotton in the middle-and-lower Yangtze River, China. Agric. Water. Manag. 283: 108313. https://doi.org/10.1016/j.agwat.2023.108313

Saeed, M., M. Naeem, A. Javed, S. Perveen, I. Sajjad, M.Z. Yousaf, M.S.M. Chohan, M. Riaz, S. Ullah, X.L. Song (2024). Characterization of water-deficit tolerance in Upland cotton (*Gossypium hirsutum* L.) assessing morphological, biochemical, molecular and yield attributes. Acta. Physiol. Plant. 46(1):13. https://doi.org/10.1007/s11738-023-03641-5

Saeed, M., X.L. Song, X.Z. Sun and M. Riaz (2018). Leaf cuticular wax content is involved in cotton leaf curl virus disease resistance in cotton (*Gossypium hirsutum* L.). Spanish. J. Agric. Res. 16(4): e0705. https://doi.org/10.5424/sjar/2018164-13085

Shah, W.U.H., G. Hao, R. Yasmeen, H. Yan and Y. Qi (2024). Impact of agricultural technological innovation on total-factor agricultural water usage efficiency: Evidence from 31 Chinese Provinces. Agric. Water. Manag. 299: 108905. https://doi.org/10.1016/j.agwat.2024.108905

Shahnawaz, M., W. Malik, M.Q. Ahmad and S. Ul-Allah (2025). Characterization and association analyses of fiber quality and biochemical traits of cotton (*Gossypium hirsutum*) under varying levels of bio-stimulant. Int. J. Agric. Biol. 33: 2025. https://doi.org/10.17957/IJAB/15.2314

Shani, M.Y., A. Ditta and M.K.R. Khan (2025). Characterizing drought-resilient cotton genotypes through morpho-physiological and biochemical traits at flowering stage. BMC. Plant. Biol. 25: 1420. https://doi.org/10.1186/s12870-025-07078-5

Shea, Z. (2023). Enhancing seed quality and application of soybean through high test weigh and low trypsin inhibitors. PhD dissertation thesis. Virginia Tech, VA 24061, USA.

Shi, X., C. Feng, H. Qin, J. Wang, Q. Zhao, C. Jiao and Y. Zhang (2024) Identification of QTNs and their candidate genes for boll number and boll weight in upland cotton. Genes. 15(8):1032. https://doi.org/10.3390/genes15081032

Shui, G., H. Lin, X. Ma, B. Zhu, P. Han, N. Aini, C. Guo, Y. Wu, Z. Pan, C. You, G. Song and X. Nie (2024). Identification of SSR markers linked to the abscission of cotton boll traits and mining germplasm in cotton. J. Cotton. Res. 7: 20. https://doi.org/10.1186/s42397-024-00180-3

Simkin, A.J., L. Kapoor, C.G.P. Doss, T.A. Hofmann, T. Lawson and S. Ramamoorthy (2022). The role of photosynthesis related pigments in light harvesting, photoprotection and enhancement of photosynthetic yield in planta. Photosynth. Res. 152(1): 23-42. https://doi.org/10.1007/s11120-021-00892-6

Stirbet, A., D. Lazár, Y. Guo, G. Govindjee (2020). Photosynthesis: basics, history and modelling. Ann. Bot. 126(4): 511-537. https://doi.org/10.1093/aob/mcz171

Sun, F., J. Ma, W. Shi and Y. Yang (2023). Genome-wide association analysis revealed genetic variation and candidate genes associated with the yield traits of upland cotton under drought conditions. Front. Plant. Sci. 14: 1135302. https://doi.org/10.3389/fpls.2023.1135302

Ul-allah, S., T. Sugumar, S.A. Jahin, B. Bhandari and H. Zhang (2026). Genetic and biotechnological innovations for drought and heat resilient cotton production: a review. J. Cotton. Res. 9: 21. https://doi.org/10.1186/s42397-026-00266-0

Van Deynze, A.E., F.J. Sundstrom and K.J. Bradford (2005). Pollen-mediated gene flow in California cotton depends on pollinator activity. Crop. Sci. 45: 1565-1570. https://doi.org/10.2135/cropsci2004.0463

Wang C.Y., A. Isoda, M.S. Li and D.L. Wang (2007). Growth and eco-physiological performance of cotton under water stress conditions. Agric. Sci. China. 6(8): 949–955. https://doi.org/10.1016/S1671-2927(07)60133-3

Ye, Y., P. Wang, M. Zhang, M. Abbas, J. Zhang, C. Liang, Y. Wang, Y. Wei, Z. Meng and R. Zhang (2023). UAV-based time-series phenotyping reveals the genetic basis of plant height in upland cotton. Plant. J. 115: 937-951. https://doi.org/10.1111/tpj.16272

Yunus, M.H., A. Firdaus, Z. Khan and M.Y.K. Ansari (2025). Genomics-assisted breeding (GAB) for trait improvement: Unveiling genomic strategies for accelerated crop enhancement. In: Zeba, K. (Ed.), Plant Breeding Technology:

Future Trends and Challenges, CABI, Oxfordshire (UK), pp. 138-165. https://doi.org/10.1079/9781800626638.0007

Zahid, K.R., F. Ali, F. Shah, M. Younas, T. Shah, D. Shahwar, W. Hassan, Z. Ahmad, C. Qi, Y. Lu, A. Iqbal and W. Wu (2016). Response and tolerance mechanism of cotton *Gossypium hirsutum* L. to elevated temperature stress: A Review. Front. Plant. Sci. 7: 937. https://doi.org/10.3389/fpls.2016.00937

Zahid, Z., M.K.R. Khan, A. Hameed, M. Akhtar, A. Ditta, H.M. Hassan and G. Farid (2021). Dissection of drought tolerance in upland cotton through morpho-physiological and biochemical traits at seedling stage. Front. Plant. Sci. 12: 627107. https://doi.org/10.3389/fpls.2021.627107

Zhao, B., Z. Liu, S.T. Ata-Ul-Karim, J. Xiao, Z. Liu, A. Qi, D. Ning, J. Nan and A. Duan (2016). Rapid and nondestructive estimation of the nitrogen nutrition index in winter barley using chlorophyll measurements. Field. Crops. Res. 185: 59–68. https://doi.org/10.1016/j.fcr.2015.10.021

Zhu, G., S. Hou, X. Song, W. Xing, W. Wei, Q. Chen and W. Guo (2021). Genome-wide association analysis reveals quantitative trait loci and candidate genes involved in yield components under multiple field environments in cotton (*Gossypium hirsutum*). BMC. Plant. Biol. 21: 250. https://doi.org/10.1186/s12870-021-03009-2