Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Genetic dissection of fiber quality traits in an interspecific Gossypium barbadense × Gossypium darwinii recombinant inbred lines population

Muhammad Iyad ABDULLAHI
State key laboratory of Cotton Bio-breeding and integrated Utilization, Institute of Cotton Research, Chinese Academy of Agricultural Sciences (ICR CAAS)

Ferehewoit Deressegn LAKEW
State key laboratory of Cotton Bio-breeding and integrated Utilization, Institute of Cotton Research, Chinese Academy of Agricultural Sciences (ICR CAAS)

Isah Mansur AMINU
State key laboratory of Cotton Bio-breeding and integrated Utilization, Institute of Cotton Research, Chinese Academy of Agricultural Sciences (ICR CAAS)

Mayamiko MASANGANO
State key laboratory of Cotton Bio-breeding and integrated Utilization, Institute of Cotton Research, Chinese Academy of Agricultural Sciences (ICR CAAS)

Asmamaw MEKONEN AMOGNE
State key laboratory of Cotton Bio-breeding and integrated Utilization, Institute of Cotton Research, Chinese Academy of Agricultural Sciences (ICR CAAS)

Bedane GUDETA
State key laboratory of Cotton Bio-breeding and integrated Utilization, Institute of Cotton Research, Chinese Academy of Agricultural Sciences (ICR CAAS)

Krisna CHANDRA SAHA
State key laboratory of Cotton Bio-breeding and integrated Utilization, Institute of Cotton Research, Chinese Academy of Agricultural Sciences (ICR CAAS)

Nuradeen MUSA
Ahmadu Bello University

Beenish NAEEM
State key laboratory of Cotton Bio-breeding and integrated Utilization, Institute of Cotton Research, Chinese Academy of Agricultural Sciences (ICR CAAS)

Waqas MUHAMMAD
State key laboratory of Cotton Bio-breeding and integrated Utilization, Institute of Cotton Research, Chinese Academy of Agricultural Sciences (ICR CAAS)

Zhongli ZHOU

State key laboratory of Cotton Bio-breeding and integrated Utilization, Institute of Cotton Research, Chinese Academy of Agricultural Sciences (ICR CAAS)

Xiaoyan CAI

State key laboratory of Cotton Bio-breeding and integrated Utilization, Institute of Cotton Research, Chinese Academy of Agricultural Sciences (ICR CAAS)

Jie ZHENG

State key laboratory of Cotton Bio-breeding and integrated Utilization, Institute of Cotton Research, Chinese Academy of Agricultural Sciences (ICR CAAS)

Fang LIU

liufcri@163.com

State key laboratory of Cotton Bio-breeding and integrated Utilization, Institute of Cotton Research, Chinese Academy of Agricultural Sciences (ICR CAAS)

Yuqing HOU

State key laboratory of Cotton Bio-breeding and integrated Utilization, Institute of Cotton Research, Chinese Academy of Agricultural Sciences (ICR CAAS)

Yanchao XU

State key laboratory of Cotton Bio-breeding and integrated Utilization, Institute of Cotton Research, Chinese Academy of Agricultural Sciences (ICR CAAS)

Research Article

Keywords: Fiber quality traits, RIL Population, QTL Mapping, Genetic Map Construction, Correlation, Expression Analysis, ANOVA

Posted Date: November 2nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11037426/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Genetic dissection of fiber quality traits in an interspecific *Gossypium barbadense* × *Gossypium darwinii* recombinant inbred lines population

ABDULLAHI Muhammad Iyad[1,2], LAKEW Ferehewoit Deressegn [1,3], AMINU Isah Mansur [1,5], MASANGANO Mayamiko [1], AMOGNE Asmamaw Mekonen[1], GUDETA Bedane[1], CHANDRA Saha Krisna[1], MUSA Nuradeen [2], NAEEM Beenish [1], MUHAMMAD Waqas[1], ZHOU Zhongli [1], CAI Xiaoyan [1,4], ZHENG Jie[1,4], LIU Fang [1,4]*, HOU Yuqing [1]*, XU Yanchao [1,4]*

[1] State Key Laboratory of Cotton Bio-breeding and Integrated Utilization, Institute of Cotton Research, Chinese Academy of Agricultural Sciences (ICR, CAAS), Anyang, Henan 455000, China.
[2] Ahmadu Bello University, P.M.B. 1044, Samaru, Zaria, Kaduna State, Nigeria.
[3] College of Agriculture and Natural Resources, Assosa University, Assosa, Ethiopia.
[4] National Nanfan Research Institute, Chinese Academy of Agricultural Sciences, Sanya, Hainan 572024, China.
[5] Department of Biological Science, Northwest University, Muhammadu Buhari Way, Kano P.M.B. 3220, Kano State, Nigeria
***Corresponding authors**: liufcri@163.com, houyuqing18@163.com, and xuyanchao2016@163.com

Abstract

Cotton (*Gossypium spp.*) is one of the most important fiber crops worldwide, providing natural fibers that support the textile industry and human livelihoods. Fiber quality traits are major determinants of cotton economic value because they directly influence spinning performance, yarn strength, and the quality of textile products. However, the genetic mechanisms underlying the complex inheritance of fiber quality traits remain incompletely understood, particularly in populations derived from wild and cultivated cotton relatives.
In this study, a recombinant inbred line (RIL) population developed from an interspecific cross between *Gossypium barbadense* and *Gossypium darwinii* was used to dissect the genetic basis of fiber quality traits. A high-density genetic linkage map was constructed using single nucleotide polymorphism (SNP) markers, comprising 3 518 markers spanning a total genetic distance of 3 801.17 cM. Phenotypic evaluations of three yield-related traits and five fiber quality traits were conducted across three consecutive years. A total of 89 quantitative trait loci (QTLs) were identified, including 20 stable QTLs consistently detected across multiple environments. These QTLs were associated with boll weight (BW), fiber length (FL), fiber strength (FS), fiber elongation (FE), seed index (SI), micronaire (MIC), and lint percentage (LP). Major QTLs were distributed mainly on chromosomes chr-D07, chr-A12, chr-D01, chr-D06, chr-D04, and chr-D12.

A QTL hotspot region on chr-D07 was identified, where co-localized loci controlling FL, FS, FE, and SI were detected. Within this hotspot, two candidate genes, *ZINC FINGER PROTEIN* and *LOC107929956*, were identified as potential regulators of fiber development. Among them, *ZINC FINGER PROTEIN* showed significantly higher expression in developing fiber and ovule tissues at 10, 15, 20, and 25 days post-anthesis (DPA), indicating its potential role in fiber formation. This study provides new insights into the genetic architecture of cotton fiber yield and quality traits by identifying stable QTLs and promising candidate genes derived from *G. barbadense* and *G. darwinii*. The identified QTL hotspot and candidate genes represent valuable genetic resources for marker-assisted selection and molecular breeding strategies targeted at improving fiber quality while maintaining high yield potential in cotton.

Keywords: Fiber quality traits; RIL Population; QTL Mapping; Genetic Map Construction; Correlation; Expression Analysis; ANOVA.

Abbreviations

BW = Boll weight

FL = Fiber length

FS = Fiber strength

FE = Fiber elongation

MIC = Micronaire

LP = Lint percentage

SI = Seed index

FU = Fiber uniformity

HVI = High volume instrument

QTL = Quantitative trait loci

RIL = Recombinant inbred line

DPA = Days post-anthesis

SNP = Single nucleotide polymorphism

MAS = Marker-assisted selection

BWA = Burrows-Wheeler alignment

GATK = Genome Alignment Tool Kits

ICIM = Interval composite interval mapping

1. Introduction

Cotton (*Gossypium* spp.) is among the world's most economically important crops, with its fiber serving as the primary natural raw material for the textile industry. Advances in textile technology, coupled with increasing consumer demand for high-quality textile products, have resulted in greater requirements for improved cotton fiber quality (Zhu and Weir, 1998). The quality of cotton fiber is a major determinant of its commercial value, as it influences the spinning and weaving processes that determine the final texture and durability of textile products. Key fiber quality traits such as fiber length, strength, fineness, and uniformity are critical for high-end textile manufacturing. The quality of cotton fiber is largely determined by its inherent physical properties, including fiber length, strength, and fineness (Zumba and Rodgers, 2016). Among these characteristics, fiber length is regarded as one of the most critical quality parameters, as it directly influences cotton value in raw material markets and plays a key role in yarn production and processing performance (Kuang and Yu, 2015). Fiber strength and micronaire (a measure of fiber fineness and maturity) have become increasingly important parameters for determining cotton bale value in the global market. Fiber strength reflects the tensile strength and durability of individual fibers, whereas micronaire provides an indication of both fiber fineness (linear density) and the degree of cell wall development and maturity during fiber formation (Zumba and Rodgers, 2016). Cotton fiber production faces increasing competition from synthetic fibers, which have gained a dominant position in the global textile market. Synthetic fibers currently account for more than

75% of global textile fiber consumption, posing a significant challenge to the competitiveness and demand for natural cotton fibers (ICAC, 2021). Therefore, the increasing competition from synthetic fibers has intensified the demand for cotton varieties with superior fiber quality and enhanced spinning performance.

However, genetic improvement of cotton fiber quality is constrained by the limited genetic diversity within upland cotton (*Gossypium hirsutum*). Consequently, expanding the genetic base through the introgression of diverse germplasm resources is essential for developing new cotton varieties with improved fiber quality and high yield potential (Abdurakhmonov et al. 2008, 2009; Yu et al. 2014). Creating new varieties using traditional breeding methods, specifically, the introduction of genes of desirable traits to the elite cotton from a donor source into the elite, is very laborious and requires at least 10 years of hard work (Abdurakhmonov et al. 2009). Simultaneous improvement of multiple desirable traits in cotton remains challenging because unfavorable correlations often exist among important characteristics, such as fiber quality, fiber yield, and tolerance to abiotic stresses. These antagonistic relationships can limit the efficiency of selection and complicate breeding efforts aimed at achieving balanced genetic improvement (Griffith and Crowfoot, 1934; Muthukumaran, 2016; Nicholson, 1960). The integration of DNA marker technologies into plant breeding has provided powerful tools for characterizing genetic diversity, identifying genomic regions associated with important traits, and accelerating the improvement of economically valuable crop characteristics (Preetha and Raveendren, 2008). The development of molecular marker technologies has provided breeders with efficient approaches for dissecting complex quantitative traits and identifying genomic regions underlying important agronomic characteristics. These technologies enable more precise selection, reduce the time and resources required for the development of new varieties and hybrids, and complement conventional breeding approaches. In recent years, numerous DNA markers and genes associated with resistance to biotic and abiotic stresses, yield-related traits, and quality attributes have been identified and mapped across diverse crop species, facilitating the application of genomics-assisted breeding (Zhang et al. 2013; Han et al., 2019; Wang et al. 2019). Molecular markers have significantly enhanced the efficiency, precision, and speed of crop improvement by facilitating the identification and selection of favorable alleles through marker-assisted selection (MAS). By enabling the early selection of plants carrying desirable genetic variants, MAS reduces reliance on phenotypic evaluation and

minimizes the influence of environmental variation on trait assessment. Recent advancements in molecular breeding approaches, including QTL mapping, genomics, high-throughput genotyping, genomic selection, and genome editing, have further expanded the application of DNA markers in modern crop improvement programs. Consequently, MAS has become an essential component of genomics-assisted breeding strategies aimed at accelerating the development of improved crop varieties (Hasan et al. 2021; Zhang et al. 2023; Tyagi et al. 2024).

Among the various cotton species, *Gossypium barbadense* (Sea Island cotton) is well-known for its exceptional fiber quality, particularly its extra-long staple (ELS) fibers, which are sought after for producing fine and strong yarns. Its superior fiber quality makes it a preferred parent in breeding programs aimed at improving fiber length, strength, and fineness. Despite its excellent fiber quality, *G. barbadense* is susceptible to biotic and abiotic stresses, which can limit its performance in certain growing conditions. In contrast, *Gossypium darwinii*, a wild cotton species native to the coastal areas of Australia, is known for its high resilience to abiotic stresses such as drought and biotic stresses including Fusarium and Verticillium wilt resistance. Although *Gossypium darwinii* generally exhibits shorter fiber length than *G. barbadense*, its rich genetic diversity and enhanced stress tolerance provide valuable genetic resources for improving cotton adaptability to environmental changes and strengthening crop resilience under future climatic challenges (Zhang et al. 2020). The hybridization of *G. barbadense* and *G. darwinii* crosses provides an excellent opportunity for the identification of QTLs that influence both fiber quality traits and biotic and abiotic stress tolerance. This crossbreeding approach is particularly useful for developing cotton varieties that are well-suited to harsh environments while maintaining desirable fiber traits. Through QTL mapping in RIL populations derived from such crosses, researchers can identify loci that influence fiber strength, fiber length, micronaire, and fiber uniformity while simultaneously introducing disease resistance and drought tolerance traits from *G. darwinii.* RIL populations are particularly valuable in QTL mapping because they are genetically stable and homozygous, which means that they provide reliable and consistent data for identifying QTLs.

Cotton fiber quality traits are governed by complex genetic networks and are strongly influenced by environmental conditions. As quantitative and polygenic traits, fiber characteristics are regulated by numerous genes with intricate interactions among themselves and with environmental factors. Therefore, improving cotton fiber quality requires a comprehensive understanding of the

underlying genetic architecture, which can be achieved through approaches such as genome-wide association studies (GWAS) and quantitative trait locus (QTL) mapping. These genomic approaches enable the identification of specific chromosomal regions and candidate genes associated with key fiber attributes, including fiber length, strength, fineness, and uniformity, thereby providing valuable genetic resources for cotton improvement programs (Wang et al. 2016; Zhang et al. 2019). The application of molecular markers in QTL mapping enables breeders to identify and monitor genomic regions associated with target traits in breeding populations, thereby improving selection efficiency and accelerating the development of superior cotton varieties. In cotton, QTL mapping has successfully identified numerous loci associated with fiber quality traits, including major-effect QTLs that contribute significantly to fiber strength and length (Li et al. 2017). These QTLs are distributed across multiple chromosomes, with individual loci generally explaining only a limited proportion of the phenotypic variation. However, the pyramiding of favorable alleles from multiple QTLs can produce cumulative effects, leading to substantial improvements in fiber quality and providing valuable resources for molecular breeding.

The development of recombinant inbred line (RIL) populations derived from the cross between *Gossypium barbadense and Gossypium darwinii* has provided valuable genetic resources for dissecting the genetic basis of cotton fiber quality traits. Zhang et al. (2019) utilized such an RIL population to identify quantitative trait loci (QTLs) associated with fiber length and strength. Their study identified multiple fiber strength-related QTLs on chromosomes 2 and 6, which contributed to both additive and dominant genetic effects. These findings highlight the complex genetic architecture underlying fiber quality traits and demonstrate the importance of exploiting diverse germplasm resources for improving cotton fiber performance. Previous investigations have identified a large number of QTLs associated with key fiber quality traits, including fiber length, strength, uniformity, elongation, and micronaire, in populations derived from *Gossypium barbadense*. These findings highlight the importance of exotic germplasm resources for uncovering the genetic basis of fiber quality variation and facilitating the improvement of superior fiber traits in cotton breeding programs (Fan et al. 2018; Zhai et al. 2016). Chen et al. (2018) identified quantitative trait loci (QTLs) associated with fiber elongation using a *Gossypium barbadense* × *Gossypium darwinii* population and detected a major-effect QTL on chromosome 15. This QTL exhibited a significant contribution to fiber development during the elongation stage,

a critical period that largely determines final fiber length. These findings highlight the potential of utilizing wild cotton germplasm for QTL introgression into cultivated cotton to enhance fiber quality while simultaneously retaining valuable agronomic characteristics, including improved stress tolerance. In addition to fiber length and strength, QTLs for fiber uniformity have also been identified in *G. barbadense* and *G. darwinii* RIL populations.

Although *G. barbadense* is widely recognized for its superior fiber quality and *G. darwinii* represents a valuable reservoir of genetic diversity, the genetic resources and introgression potential of *G. darwinii* for improving fiber quality in *G. barbadense* remain comparatively underexplored. In light of the abovementioned, this research aims to conduct a comprehensive genetic analysis of fiber yield and quality traits in 198 RIL genotypes derived from interspecific hybridization between *G. barbadense* and *G. darwinii* cotton to identify key quantitative trait loci (QTLs) and candidate genes associated with superior fiber properties.

2. Materials and methods

2.1 Development and phenotypic evaluation of the RIL population

A cross between *G. barbadense* L. (Xinhai 18) and *G. Darwinii Wa*tt *(*Darwinii 5-7) was conducted to generate 198 RIL individuals that were used for phenotypic evaluation, genetic map construction, and QTL detection. The single-seed descent (SSD) technique was employed to develop the recombinant inbred lines (RIL). In each F2 progeny line, one individual was selected randomly and selfed to generate the next progeny line by self-hybridization. An $F_{6:7}$ population was carefully selected as the RIL population for phenotypic evaluation. Field evaluation of the RIL population was carried out for 3 consecutive growing seasons (2023, 2024, and 2025) in Xinjiang (XJ) (East 113°370. North 35°120). The experimental design was laid out on a randomized complete block design with two biological replications. Each genotype was sown on a 5 m-long single ridge with inter-row spacing of 0.8 m. Each row was planted to 15 to 20 plants per ridge. All agronomic and other cultural management practices were performed in accordance with the recommendations of the Institute of Cotton Research, CAAS (ICR, CAAS) in Anyang. Agronomic and fiber quality traits were recorded as follows: boll weight (BW) (g) measured as the weight of 30 mature bolls, lint percentage (LP) (%) calculated as the weight of 30 fibers divided by seed weight, and seed index (SI) (%) recorded as the weight of 100 cotton seeds. Fiber length (FL) (mm), fiber strength

(FS) (cN·tex^{-1}), micronaire (MIC) (μg·inch^{-1}), fiber uniformity (FU) (%), and fiber elongation (FE) (%) were measured using the HVI1000 (USTER®) machine. Descriptive statistics, heritability, variance analysis, and correlation analysis were performed using SPSS software.

2.2 Phenotypic data collection

A total of 25 naturally open bolls were hand-harvested from each plot. After ginning, 10-15 g of fiber from each sample was sent to the Cotton Fiber Quality Inspection and Testing Center of the Ministry of Agriculture, Anyang, China, and the fiber quality traits of each genotype were measured by the HVI1000 automatic fiber testing system (http://www.uster.cn/en/instruments/fiber-testing-1/uster-hvi-3/), including FE (%), FS (cN·tex^{-1}), FL (mm), MIC (μg inch^{-1}) and FU (%).

2.3 Statistical analysis

2.3.1 Analysis of variance and estimation of genetic effects

Data on yield and fibre quality traits were subjected to both separate and combined analyses of variance (ANOVAs) using SPSS. Fisher's least significant difference (LSD) test ($\alpha = 0.05$) was used to separate the means. Genotypes, replications, and environment were considered as random effects. Years were treated as environments. The following random model was used for ANOVA in one environment:

$$Y_{ijk} = \mu + G_i + \beta_j + \gamma_{k(j)} + \varepsilon_{ijk}$$ (González-Barrios et al. 2019)

Where, Y_{ijk}, μ, G_i, β_j, $\gamma_{k(j)}$ and ε_{ijk} are observations of the i^{th} genotype and j^{th} replication in the k^{th} block, overall mean, effect of the i^{th} genotype, effect of the j^{th} replication, effect of the k^{th} block nested within j^{th} replication, and random error, respectively.

The following random model was used for ANOVA across environments:

$$Y_{ijkl} = \mu + G_i + \beta_{j(l)} + \gamma_{k(lj)} + E_l + (GE)_{il} + \varepsilon_{ijkl}$$ (González-Barrios et al. 2019)

Where, Y_{ijkl}, μ, G_i, $\beta_{j(l)}$, $\gamma_{k(lj)}$, E_l, $(GE)_{il}$ and e_{ijkl} are observations of the i^{th} genotype in the j^{th} replication, the k^{th} block, and the l^{th} environment, the overall mean, effect of the i^{th} genotype, within-environment replicate effect, environment x replication within block effect, environment effect, genotype x environment interaction effect, and random error, respectively.

2.3.2 Estimation of genetic variability parameters and correlation

Broad-sense heritability as well as correlation analysis was computed using the formulae given by Singh and Chaudhary (1985) as follows:

(a) **Broad-sense heritability** $(h_b^2) = \frac{\sigma_g^2}{\sigma_{ph}^2}$

Where, $\sigma_{ph}^2 = \sigma_g^2 + \frac{\sigma_{gl}^2}{l} + \frac{\sigma_e^2}{rl}$ as described according to (Littell et al. 2006), σ_{ph}^2, σ_g^2, r and l are phenotypic variance, genotypic variance, number of replications, and number of environments, respectively. Heritability estimates were grouped as high (> 50 %), moderate (20 – 50 %), and low (< 20 %), as suggested by Stansfield (1986).

(b) **Correlation coefficients** $(r_{xy}) = \frac{Cov(x,y)}{\sqrt{\sigma^2{}_x \times \sigma^2{}_y}}$

Where, $Cov(x,y)$ $\sigma^2{}_x$, $\sigma^2{}_y$ are the covariance between the traits x and y, the variance of trait x, and the variance of trait y, respectively.

2.4 SNP Marker Calling

The two parental lines and 198 RILs were used for marker calling. The samples' genomes were randomly cut using the Illumina HiSeq2500 platform. 380 bp to 480 bp genome segments were obtained for further sequencing. The raw data were filtered according to published methods by Wallace and Mitchell (2017). The Burrows-Wheeler Alignment (BWA) program was used to align the purified data to the reference genome (https://cottonfgd.net/about/download/assembly/genome.Gbar.CRI.fa.gz), while SNPs were identified using Genome Alignment Tool Kits (GATK) and Sequence Alignment Tools (SAM). The polymorphic SNP markers were used to construct a genetic map.

2.5 Genetic Map Construction

Genetic map construction was performed using the genotype-by-sequencing method, which involves identification of SNP markers. Markers that covered < 60 % of the progeny individuals in the RIL were discarded. Segregation of each SNP marker was conducted using a Chi-square test. Markers with distorted inheritance patterns were used to balance marker distribution across each

chromosome and mitigate segregation distortion. The filtered SNP markers were employed to construct a genetic map using the Ici-Mapping software (V. 4.2).

2.6 QTL Mapping

Ici-Mapping software version 4.2 was used to detect QTLs associated with all the traits under consideration. Interval composite interval mapping (ICIM) parameters were used to map the QTLs. Forward and backward regression models with a 1 cM walking speed, a probability of 0.01 for input and output, and a window size of 10 cM were used. The LOD threshold was determined by the manual permutation technique, establishing a value of 11.5, corresponding to 2.5 LOD scores. Trait QTL values below 2.5 scores were considered not fit for further selection of significant and consistent QTLs.

2.7 Analysis of Loci and Selection of Candidate Gene

To comprehensively capture potential functional genes and cis-regulatory elements linked to target traits, on chromosome D07, the original bin interval (Bin_1605594_1613130) was extended by 10 kb upstream and downstream. The information on candidate genes in this QTL region was extracted directly from the *Gossypium barbadense* reference genome accessible at the CottonGen website (https://www.cottongen.org). Additionally, interspecific orthology analysis was performed to identify their one-to-one orthologs in *G. hirsutum*.

3. Results

3.1 Phenotypic Analysis and Correlation of Traits in RIL Population of *G. barbadense* and *G. darwinii* Introgressions

The phenotypic evaluation was conducted in a recombinant inbred line (RIL) population comprising 198 individuals and two parental lines, Xinhai-18 and Darwinii 5-7. Results from the phenotypic data analysis showed abundant variation among accessions. BW, LP, FL, FU, FS, MIC, FE and SI exhibited values in the ranges of 2.28-4.93 g, 20.40-39.57 %, 25.87-38.77 mm, 82.13-88.90 %, 27.43-47.10 $cN \cdot tex^{-1}$, 2.53-5.63 μg $inch^{-1}$, 6.57-7.20 % and 9.8-16.45 % with average values of 3.5 g, 31.57 %, 33.14 mm, 86.24 %, 37.11 $cN \cdot tex^{-1}$, 4.45 %, 6.94 % and 12.75 %, respectively (Table 1). All yield and fiber quality traits were normally distributed (Fig. 1), indicating that these traits are quantitative and controlled by multiple genes. The coefficients of

variation (CVs) for BW, LP, FL, FU, FS, MIC, FE and SI were 12.79 %, 10.19 %, 8.36 %, 1.47 %, 11.16 %, 13.04 %, 1.89 % and 9.38 %, respectively (Table 1). Correlation analysis demonstrated significant associations among the majority of the evaluated traits (Fig. 2). BW exhibited highly significant positive correlations with all other measured traits, indicating a strong relationship between boll productivity and fiber-related characteristics. LP showed significant positive correlations with FE, MIC, and FS. FL was positively correlated with SI, FE, FS, and FU. Similarly, FU displayed significant positive associations with SI, FE, MIC, and FS. FS was positively correlated with SI, FE, and MIC. MIC showed a significant positive correlation with FE, and FE was positively associated with SI. In contrast, a significant negative correlation was observed between LP and SI. No significant correlations were detected among the remaining trait combinations, indicating limited associations between these trait pairs. In addition, ANOVA revealed that the effects of genotype (G), year (Y), and the interaction of genotype and year (G × Y) were significant (P<0.001) for all traits (Table S1). This result suggested that fiber quality traits are affected by both genotype and year. The broad-sense heritability (H^2) values for BW, LP, FL, FU, FS, MIC, FE, and SI were 90.20 %, 94.97 %, 95.00 %, 71.9 %, 93.25 %, 93.45 %, 75.62 %, and 93.4 %, respectively (Table 1).

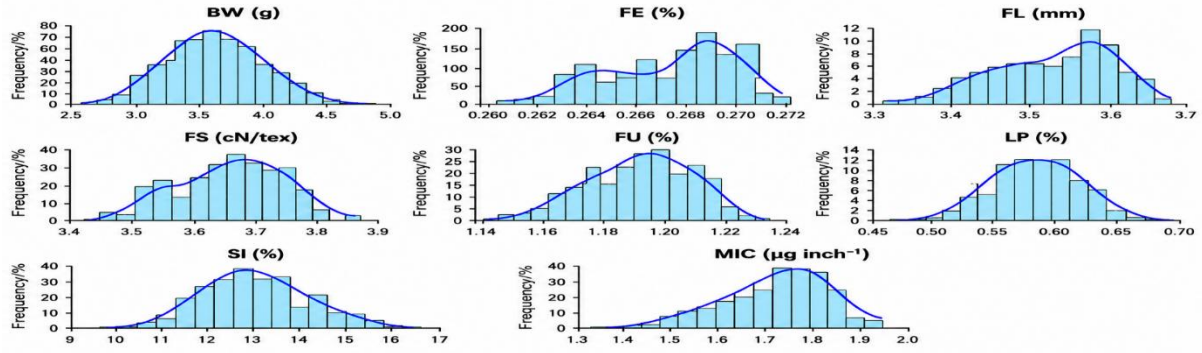


Fig. 1 Frequency distributions of yield and fiber quality traits in a RIL population of a cross between *G. barbadense* and *G. darwinii* cotton. Boll weight (BW), fiber elongation (FE), fiber micronaire (MIC), fiber strength (FS), fiber length (FL), fiber uniformity (FU), Seed index (SI), and lint percentage (LP).

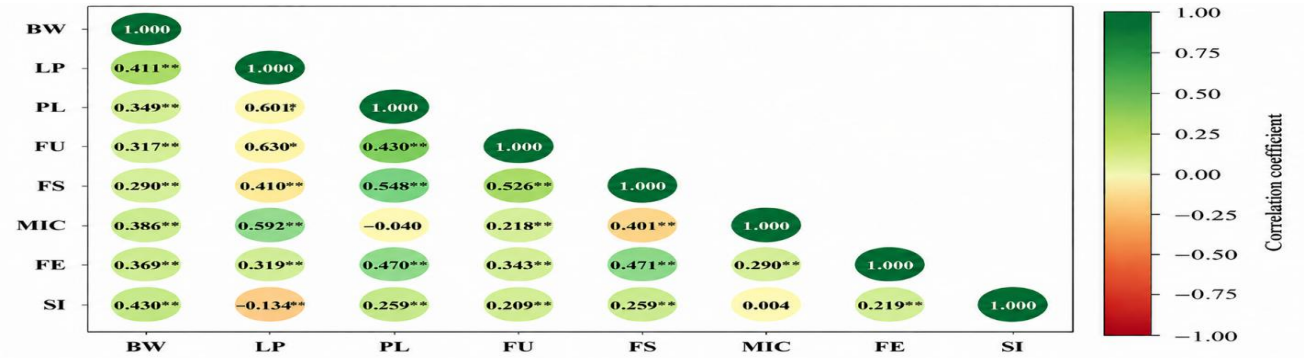


Figure 2 Correlation analysis of yield and fiber quality traits in the a RIL population of a cross between *G. barbadense* and *G. darwinii* cotton. Boll weight (BW), fiber elongation (FE), fiber micronaire (MIC), fiber strength (FS), fiber length (PL) and fiber uniformity (FU), Seed index (SI), Lint percentage (LP). * indicates significant correlation at P<0.05; ** indicates significant correlation at P <0.01.

Table 1. Phenotypic statistics of yield and fiber quality traits of a RIL population between *G. barbadense* and *G. darwinii Cotton*

Traits	Min	Max	Mean	SD	CV (%)	H^2(%)
BW	2.28	4.93	3.51	0.44	12.79	90.20
LP	20.40	39.57	31.57	3.21	10.19	94.97
FL	25.87	38.77	33.14	2.77	8.36	95.00
FU	82.13	88.90	86.24	1.26	1.47	71.69
FS	27.43	47.10	37.11	4.14	11.16	93.25
MIK	2.53	5.63	4.45	0.57	13.04	93.94
FE	6.57	7.20	6.94	0.13	1.89	75.62
SI	9.84	16.45	12.75	1.19	9.38	93.24

Note: BW = Boll weight; LP = Lint percentage; FL = Fiber length; FU = Fiber uniformity; FS = Fiber strength; MIC = Micronaire; FE = Fiber elongation; SI = Seed index; Min = Minimum; Max = Maximum; SD = Standard deviation; CV = Coefficient of variation; H^2 = Broad sense heritability.

3.2 Genotypic and SNP Marker Analysis in a Recombinant Inbred Line Population.

A total of 200 individuals, including 198 RILs and two parental lines, were subjected to genome-wide resequencing analysis. The average number of generated reads per individual was 183.91 million, of which 183.33 million were uniquely aligned to the reference genome**.** The average mapping rate reached 99.67 %**,** while only 0.58 million reads per sample remained unmapped, demonstrating the high reliability and accuracy of the sequencing dataset for downstream SNP discovery and genetic analysis (Additional Table 2). After filtering and quality control, the total number of detected SNPs across all individuals was 797 239 111, consisting of approximately 551 480 141 transitions and 245 758 970 transversions. On average, 29.16 % of these markers showed heterozygosity (Additional Table 3). The SNP markers were further refined by removing markers that did not align with the average sequencing depth of the parental lines, which was set at 8x. Ultimately, 3 518 SNP markers were selected for construction of the genetic map.

3.3 Construction and Statistical Description of Genetic Map

A total of 3 518 polymorphic homozygous SNP markers between the two parents were used to estimate recombination frequency and genetic distance using QTL IciMapping v4.2. Based on genetic distance, SNP markers were assigned to 26 linkage groups corresponding to the 26 chromosomes of *G. barbadense* and renamed according to the reference genome physical map. (Figure 3). The genetic distance between adjacent markers was estimated, and finally the total genetic map distance was 3 801 cM, with 2 091.7 cM in the A sub-genome and 1 709.5 cM in the

D sub-genome. The average number of markers in each linkage group was 135.3; chromosome A04 contained the fewest markers (10), while chromosome D22 contained the most (270markers). The average distance between the 26 linkage groups was 146.2 cM; chromosome A04 showed the shortest distance at 2.64 cM, while chromosome A05 showed the longest at 328.98 cM. The largest gap observed on chromosome D23 was 38.98 cM. The average proportion of gaps less than 5 cM across the 26 linkage groups was 96.71% (Table 2).

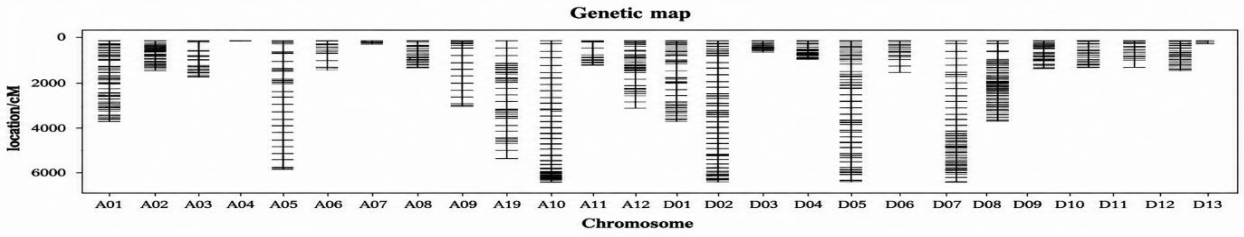


Figure 3. Shows the genetic map of 198 RIL population derived from a cross between *G. bar* × *G. darwinii*. The X-axis represents the number of chromosomes, where A01 to A13 indicate the 13 chromosomes of the A-subgenome of *G. barbadense* and D01 to D13 indicate the 13 chromosomes of the D-subgenome of *G. darwinii*. The Y-axis represents the genetic distance (cM, centiMorgan), and the black lines across the chromosomes indicate the markers.

Table 2. Detailed information of SNP markers in the genetic map

Chromosomes	Total Bin Markers	Total Distance (cM)	Average Distance(cM)	Maximum Gap(cM)	Gaps<5 (cM)%
A1	120	215.28	1.79	28.2	92.50%
A2	104	56.62	0.54	4.5	100%
A3	191	155.25	0.81	8.7	97.90%
A4	10	2.64	0.26	0.45	100%
A5	165	328.98	1.99	28.77	89.70%
A6	67	89.01	1.33	16.17	94%
A7	47	101.03	2.15	26.63	93.60%
A8	69	44.87	0.65	4.17	100%
A9	126	93.51	0.74	4.93	100%
A10	156	196.83	1.26	36.24	95.50%
A11	196	140.84	0.72	8	97.95%
A12	155	198.13	1.28	17.62	94.84%
A13	205	215.70	1.05	14.8	96.59%
D1	159	252.35	1.59	29.95	94.96%
D2	155	177.37	1.14	29.59	95.48%
D3	56	54.60	0.98	7.24	96.42%
D4	79	78.80	1.00	8.52	94.94%
D5	226	291.14	1.29	24.82	96%
D6	139	121.40	0.87	34.66	97.84%
D7	163	138.23	0.85	9.88	98.15%
D8	190	178.25	0.94	16.82	95.26%

D9	270	206.69	0.77	12.17	97.41%
D10	112	139.53	1.25	38.98	97.32%
D11	121	125.07	1.03	20.87	97.52%
D12	160	117.49	0.73	6.1	98.75%
D13	77	81.58	1.06	8.99	96.05%
Total	3518	3801.21	1.07	17.22	96.49%

Note: cM = CentiMorgan

3.4 QTL Analysis and Allelic Effect of High-Density Genetic Map

According to the QTL study, 89 QTLs were detected for all the traits, with 20 and 69 QTLs distributed across A and D subgenomes, respectively. Specifically, there were 12 QTLs for BW, 15 for SI, 14 for FS, 13 for FL, 10 for MIC, 13 for FE, 6 for FU, and 6 for LP. The distribution of detected QTLs varied considerably among the 18 chromosomes. Chromosome D07 harbored the highest number of QTLs (20), followed by D12 (13), D02 (11), D01 (8), and D06 (7). Moderate QTL enrichment was observed on chromosomes D04, D11, and D12, each containing five QTLs, while D05 contained 3 QTLs. Chromosomes A01, A02, and D08 each carried two QTLs, whereas chromosomes A05, A06, A07, D03, D09, and D10 each contained a single QTL. In contrast, no significant QTLs associated with the evaluated traits were detected on the remaining 8 chromosomes, indicating a non-uniform distribution of genetic loci controlling the studied traits across the cotton genome. The phenotypic variability explained (PVE) by the detected loci ranged from 1.69 to 39.78%, while the LOD scores varied from 2.50 to 31.25. (Additional Table 4).

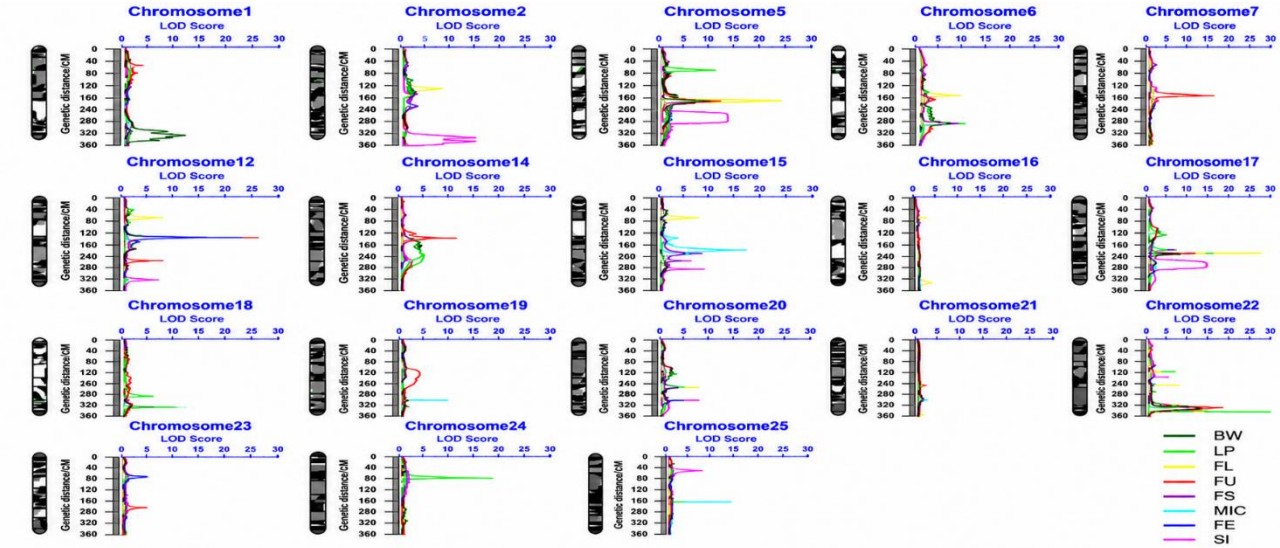


Fig. 4. QTL distribution and genomic locations associated with yield and fiber quality traits in *Gossypium barbadense*. The Y-axis represents the genetic distance (cM, centiMorgan) of each chromosome, while the X-axis indicates the chromosome number. Each colored line represents an individual QTL associated with different traits: dark green lines represent boll weight (BW), light green lines represent lint percentage (LP), yellow lines represent fiber length (FL), red lines represent fiber uniformity (FU), purple lines represent fiber strength (FS), cyan lines represent fiber micronaire (MIC), blue lines represent fiber elongation (FE), and pink lines represent seed index (SI).

3.5 QTL Clusters on Specific Chromosomes on the Genetic Map

Among the 26 chromosomes analyzed, 7 chromosomes were identified as QTL hot-spot regions, collectively harboring a high proportion of the detected QTLs associated with fiber quality and yield-related traits. Chromosome D07 represented the major QTL-rich region, containing 20 QTLs associated with FL, FS, FE, MIC, BW, and SI. Chromosome A12 carried 13 QTLs linked to BW, MIC, SI, and FS, while chromosome D01 contained 8 QTLs associated with FL, FE, FU, and LP. Furthermore, chromosome D04 harbored 5 QTLs associated with MIC, FU, and SI, chromosome D11 contained 5 QTLs related to FS, FE, and FU, and chromosome D12 carried 4 QTLs associated with FL, FS, and LP. Among the 89 QTLs identified, 20 were classified as stable QTLs because they were consistently detected across two or three years of evaluation. These stable QTLs included four QTLs associated with SI, four with FL, three with FS, three with FE, two with BW, two with MIC, one with LP, and one with FU. The stable QTLs were distributed across seven chromosomes, including D07, A12, D01, D06, D04, D11, and D12.

3.6 Analysis of Loci and Selection of Candidate Genes in the QTL regions

A total of six (6) genes were identified in the extended candidate region, including: *Gbar_D07G001580*, *Gbar_D07G001590*, *Gbar_D07G001600*, *Gbar_D07G001610*, *Gbar_D07G001620,* and *Gbar_D07G001630*. Interspecific orthology analysis revealed that each of the 6 *G. barbadense* genes had a corresponding ortholog in *G. hirsutum*, including *Ghi_D07G00731*, *Ghi_D07G00736*, *Ghi_D07G00746*, *Ghi_D07G00751*, *Ghi_D07G00756,* and *Ghi_D07G00761*. *G. barbadense* alleles predominantly enhance fiber traits, whereas *G. darwinii* alleles contribute to some fiber fineness and micronaire traits. Furthermore, correlation analysis revealed a significant positive association between Bin_1605594_1613130 and fiber quality traits (Additional Table 5B).

3.7 Genotype Phenotype Integration Analysis for Marker Bin_1605594_1613130_D07

Results from the integration analysis revealed statistically significant differences ($P < 0.001$) among the genotypes for all fiber quality traits. Xinhai 18 consistently exhibits higher mean values for fiber length, strength, elongation, uniformity, and micronaire compared with Darwin 5-7. (Additional Figure 1a).

3.8 Correlation Analysis Between Genotype and All Fiber Traits for Marker Bin_1605594_1613130_D07

Findings from the correlation study revealed a significant positive correlation (P<0.001) between all the fiber quality traits and the genotypes. FE shows the highest correlation with genotype (r ≈ 0.630), followed by FS (r ≈ 0.570), FL (r ≈ 0.46), MIC (r ≈ 0.410), and FU (r ≈ 0.360), all of which exhibit significant moderate correlations. (Additional Figure 1b).

3.9 Expression Analysis of Candidate Genes in the QTL Region

Expression data for the 6 candidate genes on chromosome D07 were obtained from the CottonMD database available at (https://yanglab.hzau.edu.cn/CottonMD/). According to the investigation, two genes may be significant for fiber quality characteristics, specifically: Ghi_D07G00746, a ZINC FINGER PROTEIN (ZFP) gene, demonstrated a remarkably elevated expression in both fibers and ovules at many time intervals. The second gene, Ghi_D07G00736, an uncharacterized protein LOC107929956, exhibited a modest expression level in both fibers and ovules (Fig. 5 and 6).

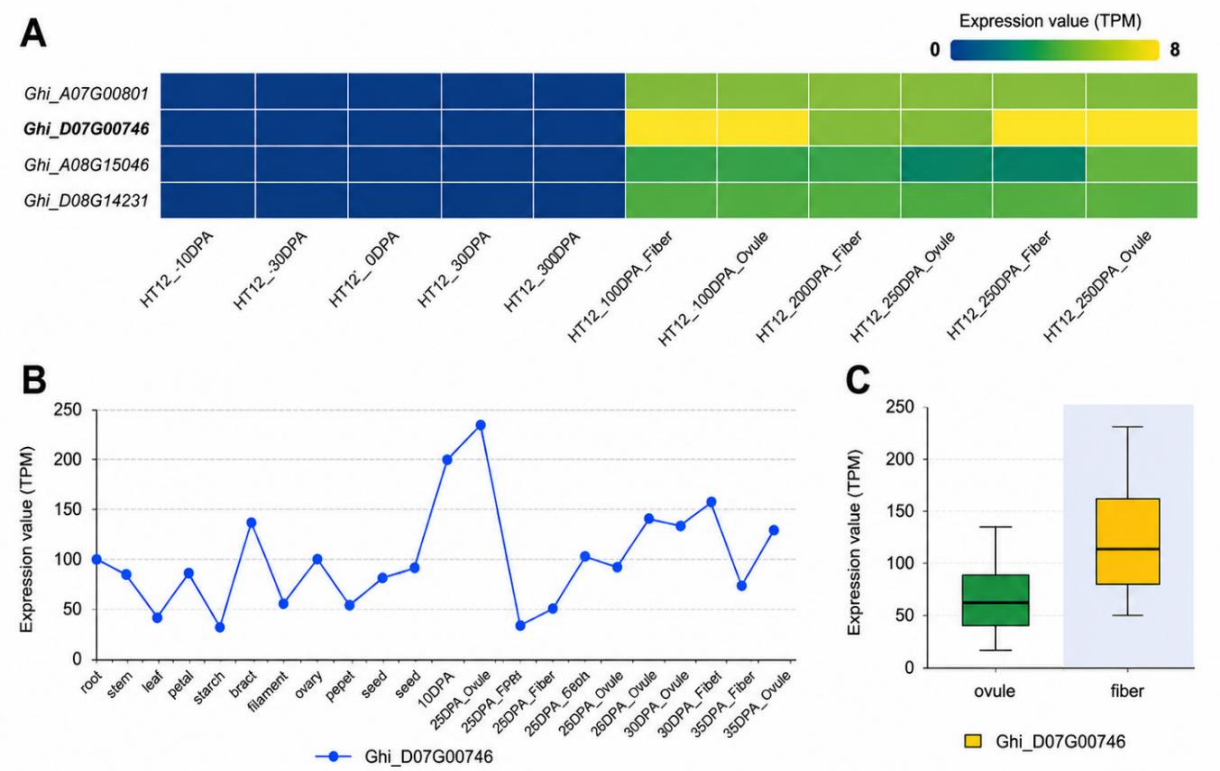


Fig. 5. Expression profiling of candidate zinc finger gene *Ghi_D07G00746* during cotton development. (A) Heat map showing expression patterns of four candidate genes across different developmental stages of cotton fibers and ovules. *Ghi_D07G00746* exhibits preferential and elevated expression in fiber and ovule tissues. (B) Tissue-specific expression profile of *Ghi_D07G00746* across various cotton tissues, including vegetative organs, floral organs, seeds, and fibers at different developmental stages. (C) Developmental expression analysis of *Ghi_D07G00746* in fiber and ovule tissues. The box plot compares the expression levels of *Ghi_D07G00746* between ovule and fiber tissues, showing higher expression in fibers.

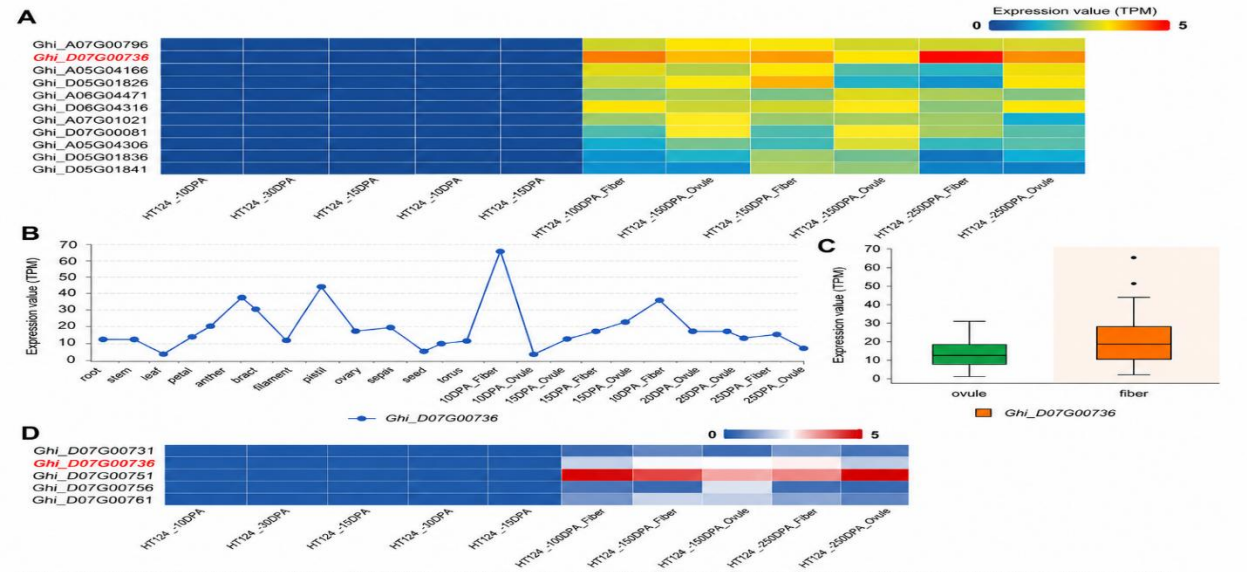


Fig. 6. Expression characterization of *Ghi_D07G00736*, candidate gene associated with cotton fiber development. (A) Heat map comparison of four candidate genes showing the preferential expression of *Ghi_D07G00736* in cotton fibers and ovules. (B) Expression pattern of *Ghi_D07G00736* across different cotton tissues. (C) Temporal expression dynamics of *Ghi_D07G00736* during ovule and fiber developent. (D) Expression heat map of six orthologous genes reveals elevated expression of *Ghi_D07G00736* in cotton fiber and ovule tissues.

4. DISCUSSION

4.1 Phenotypic Diversity and Heritability of Fiber Quality Traits

Genetic diversity plays a fundamental role in QTL mapping by providing the allelic variation required to identify associations between genomic regions and target traits. A broader genetic background increases the likelihood of detecting informative polymorphic markers and improves the precision of mapping trait-associated loci. Diverse parental germplasm enables the capture of different genetic effects, including additive, dominant, and epistatic interactions, thereby enhancing the reliability, resolution, and practical utility of identified QTLs for marker-assisted selection and crop improvement programs (Collard et al. 2005; Xu and Crouch, 2008). The considerable genetic variation detected in the present *Gossypium barbadense* × *Gossypium darwinii* RIL population corresponds with previous studies demonstrating high levels of genetic diversity within and among *Gossypium* germplasm resources, particularly between *G. barbadense* and *G. darwinii* (Ditta et al. 2018; Abdullaev et al. 2017).

To improve the reliability of the phenotypic measurements and minimize environmental variation in QTL detection, phenotypic data were collected from a single environment over three consecutive years. Phenotypic statistical results showed that all traits exhibited phenotypic variation ranging from 1.47 to 13.04 % (Table 1), which is similar to previous studies (Sun et al.

2017, Huang et al. 2017, Dong et al. 2018). Heritability is an important genetic parameter that affects the reliability and precision of association mapping by determining the proportion of phenotypic variation attributable to genetic factors (Bernardo et al. 2004). Broad-sense heritability (H^2) is commonly used to evaluate the extent to which phenotypic variation is genetically controlled and to assess the stability of inherited traits. According to Singh (2001), heritability can be classified into three categories: low (<30 %), moderate (30-70 %), and high (>70 %). In the present study, all evaluated traits exhibited high heritability, indicating a strong genetic contribution to trait variation. These findings are consistent with earlier studies that found strong heritability estimates for cotton fiber-related variables (Nie et al. 2016; Huang et al. 2017; Dong et al. 2018).

Furthermore, analysis of variance (ANOVA) revealed that the effects of genotype, year, and their interaction were significant for all evaluated traits (Additional Table 1), indicating that environmental factors contributed to the phenotypic variation of these characters. The genotype effect, which showed the greatest contribution, suggests that genetic factors predominantly govern the quality traits of cotton fibers. Therefore, these traits are suitable for genetic dissection through linkage mapping approaches. Correlation analysis among different traits is a crucial component of breeding programs, as it provides insights into the relationships and potential interactions among target characteristics. Understanding trait associations helps breeders design more efficient selection strategies and facilitates the simultaneous improvement of multiple desirable traits in cotton breeding programs (Manan et al. 2022). The correlation analysis revealed significant positive associations among all evaluated fiber quality traits, indicating that these traits may be improved simultaneously through effective selection strategies. In addition, significant positive correlations were observed between yield-related traits and fiber quality parameters, suggesting potential opportunities for concurrent improvement of productivity and fiber performance. However, lint percentage (LP) showed a slight negative correlation with seed index (SI), indicating a possible trade-off between these traits. These findings provide valuable insights for optimizing selection strategies to improve multiple fiber quality attributes in cotton breeding programs.

4.2 Construction of High-Density Genetic Map

High-density genetic maps are essential for precise QTL identification and fine mapping of genes controlling important agronomic traits. In cotton, numerous genetic maps have been developed using PCR-based molecular markers, particularly simple sequence repeat (SSR) markers. For instance, Sun et al. (2021) constructed a genetic map spanning 959.4 cM with 155 SSR markers. Also, Shang et al. (2015) developed a linkage map comprising 581 SSR markers, with an average marker interval of approximately 6.39 cM. These studies demonstrate the contribution of molecular marker-based maps to improving the resolution of genetic analysis and QTL detection in cotton. Liu et al. (2015) used 1 675 SSR markers to construct a genetic map spanning a total genetic distance of 3 338.2 cM. However, the limited number of available SSR markers has constrained the development of high-resolution genetic maps in cotton. The adoption of single-nucleotide polymorphism (SNP) markers has substantially improved marker density and enhanced the resolution of cotton linkage maps. For example, Han et al. (2022) developed a high-density genetic map using a *Gossypium barbadense* RIL population and resequencing-derived SNP markers, identifying 933 845 SNPs and constructing a linkage map spanning 2 483.17 cM. Similarly, studies in *Gossypium hirsutum* have demonstrated the effectiveness of high-throughput SNP genotyping for construction of dense genetic maps. Jia et al. (2016) developed a high-density map containing 6 295 SNP markers and 139 SSR markers**,** covering 4 071.98 cM with an average marker interval of 0.63 cM**.** Zhang et al. (2016) constructed a linkage map comprising 5 521 SNP markers, spanning 3 259.37 cM with an average marker interval of 0.78 cM. Likewise, Li et al. (2017) generated a high-density genetic map using 3 978 SNP markers, covering 2 480.36 cM with an average marker distance of 0.62 cM. Collectively, these studies highlight the advantages of SNP-based high-density mapping approaches for improving QTL resolution and facilitating the identification of genes underlying complex traits in cotton. Therefore, SNP markers are more effective for constructing genetic maps. Many of the genetic maps constructed for cotton come from *G. hirsutum.* In this study, 198 *G. barbadense* RIL populations were re-sequenced, and a high-density genetic map was constructed using 3 518 high-quality SNP markers covering a distance of 3 801.17 cM with an average of 1.0 cM. Furthermore, the genetic map identified several SNP markers located within *G. darwinii* fragments. With its extensive coverage of the entire genome of *G. barbadense* cotton, this high-density genetic map serves as a valuable tool for detecting QTLs and genes associated with both sea-island and Darwin cotton. Its large number of SNP markers enhances its utility for comprehensive genetic analyses in cotton breeding programs.

4.3 Influence of Marker Numbers, Environments, and QTL Detection Methods on QTL Identification

The accuracy and efficiency of QTL analysis are influenced by several factors, including molecular marker density, marker distribution, and the mapping strategy employed (Sabev et al. 2020). In addition, environmental variation across different growing conditions can affect QTL detection, influencing both the number of identified loci and the precision of their estimated genomic positions (Cornforth and Long, 2003). For complex quantitative traits, such as fiber yield and quality, identification of stable QTLs across multiple environments enhances their reliability and practical value for gene discovery and marker-assisted breeding programs (Kushanov et al., 2021). In the present study, the detection of stable QTLs associated with fiber yield and quality traits provides valuable genetic resources for elucidating the molecular mechanisms underlying fiber development in both *Gossypium barbadense* and *Gossypium darwinii.*

4.4 QTL Clusters and Genetic Correlation between Fiber Quality and Fiber Yield

QTL clusters harboring loci associated with multiple agronomic traits represent valuable genomic regions for the simultaneous improvement of correlated characteristics through marker-assisted selection (Collard et al. 2005; Zhang et al. 2020). These clustered QTL regions provide opportunities to pyramid favorable alleles and enhance breeding efficiency by targeting multiple desirable traits in a single selection process. Among the 26 chromosomes analyzed, QTL cluster regions were identified on seven chromosomes. Chromosome D07 harbored 20 QTLs related to BW, FL, FS, SI, FE, and MIC. Chromosome A12 contained 13 QTLs related to BW, FS, MIC, and SI. Chromosome D01 contained 8 QTLs related to FL, FE, and SI. Chromosome D06 contained 8 QTLs related to FL, FE, FU, and LP. Chromosome D04 contained 5 QTLs related to MIC, FU, and SI. Lastly, Chromosome D12 contained 4 QTLs related to FL, FS, and LP (Figure 5, Table 2). Genes and QTLs located within the same genomic regions may be physically linked or may exert pleiotropic effects, resulting in the co-localization of loci controlling multiple traits (Tan et al. 2018; Yang et al. 2022). However, QTL clusters containing favorable alleles for multiple traits provide promising genomic regions for coordinated genetic improvement and efficient selection of desirable trait combinations (Song et al. 2017; Wang et al. 2020). Consistently,

the present study suggests that genetic linkage is the primary factor underlying the difficulty in combining superior fiber quality with high-yielding traits.

4.5 Genotype-Phenotype Integration Analysis and Correlation Analysis Between Genotype and All Fiber Traits for Marker Bin_1605594_1613130_D07

Results from integration analysis revealed that the genotype distribution at marker Bin_1605594_1613130_D07 was reasonably balanced, with 105 genotypes carrying Xinhai 18, 92 genotypes carrying Darwin 5-7, and only 1 missing genotype. The very low missing rate indicated good marker quality. Comparison of fiber traits between genotypic classes showed that Xinhai 18 was consistently associated with higher fiber length, fiber strength, fiber elongation, fiber uniformity, and micronaire, suggesting that the Xinhai 18 allele at this locus is the favorable allele for fiber-quality improvement (Fig. 5a). The significant differences (P<0.001) observed between the two genotypes support their use in breeding programs for fiber improvement. Correlation analysis revealed significant positive correlations (r values ~0.350-0.630, P<0.001) between the genotypes and all the fiber quality traits (Fig. 5b). This implies that genetic variation contributes significantly to the expression of fiber quality traits in this population. Therefore, improvement in one of these traits based on genotype is likely to be associated with improvements in the others, suggesting shared genetic regulation or pleiotropy among these traits.

4.6 Expression analysis of candidate genes in candidate QTL region

Expression profiling of candidate genes within QTL regions, particularly in developing fiber and ovule tissues, provides valuable evidence prioritizing potential genes underlying fiber-quality traits and strengthening the functional interpretation of QTLs (Wang et al. 2020). The expression profiles of the identified candidate genes were investigated in ovule and fiber tissues across different stages of cotton fiber development. Among the six candidate genes located within the chromosome D07 QTL hotspot region, which harbors major and stable QTLs associated with important fiber quality traits, including fiber length (FL), fiber strength (FS), and fiber elongation (FE), two genes exhibited consistently higher transcript abundance in both ovule and fiber tissues throughout multiple developmental stages. The preferential expression patterns of these genes suggest their potential involvement in regulating ovule and fiber development and support their selection as promising candidates for further functional characterization (Fig. 6).

Among these candidates, *ZINC FINGER PROTEIN* (*ZFP*) and *LOC107929956* may contribute to the regulation of key fiber quality traits, including fiber length, strength, and elongation. In previous studies, ZFP proteins have been implicated in cotton fiber development through their roles in transcriptional regulation and ubiquitin-mediated protein degradation pathways, suggesting that these genes may influence fiber formation by modulating protein stability and developmental processes. Zinc finger proteins are important regulators of cotton fiber development. Several *GhZHD* and *GhZFP* genes show fiber-preferential expression and are involved in fiber elongation, pigment accumulation, and hormone-mediated regulation, making them promising candidates for functional validation (Abdullah et al. 2018; He et al. 2019). Following QTL identification, candidate genes can be functionally investigated using approaches such as qRT-PCR, VIGS, gene overexpression, and CRISPR/Cas9-mediated mutagenesis to clarify their roles in cotton fiber development and quality (Liu et al. 2019; Gao et al. 2017). Nevertheless, this study has some limitations. The QTLs identified here were derived from a single *G. barbadense* × *G. darwinii* RIL population evaluated across three consecutive years, and their stability across diverse genetic backgrounds and environments remains to be confirmed. In addition, the candidate genes were prioritized based on their genomic positions and expression patterns but were not functionally validated. Further studies using independent populations and functional approaches are therefore needed to confirm the causal roles of these QTLs and candidate genes in cotton fiber development.

5. **Conclusion**

In this study, a population of recombinant inbred lines (RILs) was generated by crossing *Gossypium darwinii* (Darwinii 5-7) with *Gossypium barbadense* (XH-18). Phenotypic assessments were conducted on three yield-related traits and five fiber quality traits. Using the genotypic information from both the parental lines and the RIL population, a comprehensive genetic linkage map was constructed. Subsequent QTL analysis identified 89 QTLs significantly associated with these traits. Notably, 20, 13, 8, 8, 5, and 4 QTLs were located on chromosome D07, A12, D01, D06, D04, and D12, respectively. Two genes were identified, including *ZING FINGER PROTEIN* and *LOC107929956,* which could be important for fiber quality traits in cotton. This research provides key insights into the genetics of cotton yield and quality, offering a basis for further studies and improvements in cotton breeding.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests

Funding information

This research was funded by the National Natural Science Foundation of China (32401824, 32272090), Sanya Yazhou Bay Science and Technology City (SCKJ-JYRC-2024-54), the Youth Innovation Program of the Chinese Academy of Agricultural Sciences (Y2025QC14), and the Nanfan Special Project of National Nanfan Research Institute of Chinese Academy of Agricultural Sciences (YBXM2636, YBXM2548)

Availability of data

The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request

Authors contribution

Abdullahi MI, Xu Y and Liu F designed the study. Abdullahi MI, Xu Y and Liu F conceived the study and analyzed the results. Zhou Z, Cai X, Zheng J and Hou YQ contributed to germplasm/material collection and phenotypic data collection. Xu Y and Abdullahi MI conducted the laboratory experiments, while Lakew FD assisted with the laboratory experiments. Abdullahi MI, Aminu IM and Xu Y performed the computational analyses. Abdullahi MI drafted the manuscript. Amogne AM, Gudeta B, Chandra SK, Musa N, Naeem B, Muhammad W and Masangano M contributed to manuscript revision. Abdullahi MI, Xu Y and Liu F contributed to critical revision, literature review, editing and proofreading of the manuscript. Xu Y and Liu F finalized the manuscript. All authors read and approved the final manuscript.

Authors email addresses

amuhammadiyad@gmail.com (A.M.I); frederesegn@gmail.com (L.F.D.); isahmansur86@gmail.com (A.I.M); maya.masangano@gmail.com (M.M); amogne.asmamaw@gmail.com (A.A.M); bedaneg@gmail.com (G.B); krisna.sau@gmail.com (C.S.K); musanuradeen@gmail.com (M.N); beenishnaeem1993@outlook.com (N.B); mwuaf.edu@gmailcom (M.W); zhonglizhou@163.com (Z.Z.); cxycri@163.com (C.X.); zhengjieself@163.com (Z.J); liufcri@163.com (L.F.); houyuqing18@163.com (H.Y.); xuyanchao2016@163.com (X.Y.).

Acknowledgements

We sincerely acknowledge and appreciate the support provided by the State Key Laboratory of Cotton Biology, Institute of Cotton Research (ICR), Chinese Academy of Agricultural Sciences (CAAS), Anyang, China, for providing the research platform and facilities that enabled this study. We extend our deepest gratitude to the entire research team, colleagues, friends, and all individuals who contributed directly or indirectly to the success of this work. Their valuable support, guidance, and contributions are greatly appreciated.

References

Abdullaev, A.A., Salakhutdinov, I.B., Egamberdiev, S.S., et al. (2017). Genetic diversity, linkage disequilibrium, and association mapping analyses of *Gossypium barbadense* L. germplasm. *PLoS ONE*, 12(11), e0188125. https://doi.org/10.1371/journal.pone.0188125.

Abdullah, M., Cheng, X., Cao, Y., et al. (2018). Zinc finger-homeodomain transcriptional factors (ZHDs) in upland cotton (*Gossypium hirsutum*): Genome-wide identification and expression analysis in fiber development. *Frontiers in Genetics*, 9, 357. doi:10.3389/fgene.2018.00357.

Abdurakhmonov, I. Y., Kohel, R. J., Yu, J. Z., et al. (2008). Molecular diversity and association mapping of fiber quality traits in exotic *G. hirsutum* L. germplasm. *Genomics*, 92(6), 478-487. doi:10.1016/j.ygeno. 2008.07.013.

Abdurakhmonov, I. Y., Saha, S., Jenkins, J. N., et al. (2009). Linkage disequilibrium-based association mapping of fiber quality traits in *G. hirsutum* L. variety germplasm. *Genetica*, 136(3), 401-417. doi:10.1007/s10709-008-9337-8.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1-48. doi:10.18637/jss.v067.i01.

Chen, Y., Liu, G., Ma, H., Song, Z., Zhang, C., Zhang, J., Zhang, J., Wang, F., & Zhang, J. 2018). Identification of introgressed alleles conferring high fiber quality derived from Gossypium barbadense L. in secondary mapping populations of *G. hirsutum* L. *Frontiers in Plant Science*, 9, 1023. doi:10.3389/fpls.2018.01023.

Collard, B. C. Y., Jahufer, M. Z. Z., Brouwer, J. B., and Pang, E. C. K. (2005). An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. *Euphytica*, 142, 169-196. doi:10.1007/s10681 005-1681-5.

Collard, B. C. Y., and Mackill, D. J. (2008). Marker-assisted selection: An approach for precision plant breeding in the twenty-first century. Philosophical Transactions of the Royal Society B: *Biological Sciences*, 363(1491), 557-572. doi:10.1098/rstb.2007.2170.

Cornforth, T. W., and Long, A. D. (2003). Inferences regarding the numbers and locations of QTLs under multiple-QTL models using interval mapping and composite interval mapping. *Genetical Research,* 82(2), 139-149. doi:10.1017/S0016672303006396.

Ditta, A., Zhou, Z., Cai, X., et al. (2018). Genome-wide mining and characterization of SSR markers for gene mapping and gene diversity in *Gossypium barbadense* L. and *Gossypium darwinii* G. Watt accessions. *Agronomy*, 8(9), 181. https://doi.org/10.3390/agronomy8090181.

Dong, C., Wang, J., Yu, Y., et al. (2019). Identifying functional genes influencing *Gossypium hirsutum* fiber quality. Frontiers in Plant Science, 9, 1968. doi:10.3389/fpls.2018.01968.

Fan, L., Wang, L., Wang, X., et al. (2018). A high-density genetic map of extra-long staple cotton (Gossypium barbadense) constructed using genotyping-by-sequencing based single nucleotide polymorphic markers and identification of fiber traits-related QTL in a recombinant inbred line population. *BMC Genomics*, 19, 489. doi:10.1186/s12864-018-4890-8.

Gao, W., Long, L., Tian, X., et al. (2017). Genome editing in cotton with the CRISPR/Cas9 system. *Frontiers in Plant Science*, 8, 1364. doi:10.3389/fpls.2017.01364.

González-Barrios, P., Díaz-García, L., and Gutiérrez, L. (2019). Mega-environmental design: Using genotype × environment interaction to optimize resources for cultivar testing. *Crop Science*, 59(5), 1899-1915. doi:10.2135/cropsci2018.11.0692.

Griffith, F. Ll., and Crowfoot, G. M. (1934). On the early use of cotton in the Nile Valley. *The Journal of Egyptian Archaeology*, 20(1), 5-12. doi:10.1177/030751333402000102.

Han, J., Han, D., Guo, Y., et al. (2019). QTL mapping pod dehiscence resistance in soybean (*Glycine max* L. Merr.) using specific-locus amplified fragment sequencing. *Theoretical and Applied Genetics*, 132(8), 2253 2272. doi:10.1007/s00122-019-03352-x.

Han, W., Zhao, J., Deng, X., et al. (2022). Quantitative trait locus mapping and identification of candidate genes for resistance to Fusarium wilt race 7 using a resequencing-based high-density genetic bin map in a recombinant inbred line population of *Gossypium barbadense*. *Frontiers in Plant Science*, 13, 815643. https://doi.org/10.3389/fpls.2022.815643.

Hasan, N., Choudhary, S., Naaz, N., et al. (2021). Recent advancements in molecular marker-assisted selection and applications in plant breeding programmes. *Journal of Genetic Engineering and Biotechnology,* 19, 128. DOI: 10.1186/s43141-021 00231-1.

He, P., Yang, Y., Wang, Z., et al. (2019). Comprehensive analyses of ZFP gene family and characterization of expression profiles during plant hormone response in cotton. *BMC Plant Biology*, 19, 329. doi:10.1186/s12870-019-1932-6.

Ho, M. H., Saha, S., Jenkins, J. N., and Ma, D. P. (2010). Characterization and promoter analysis of a cotton RING-type ubiquitin ligase (E3) gene. *Molecular Biotechnology*, 46(2), 140-148. doi:10.1007/s12033-010-9280-7.

Huang, C., Nie, X., Shen, C., et al. (2017). Population structure and genetic basis of the agronomic traits of upland cotton in China revealed by a genome-wide association study using high-density SNPs. *Plant Biotechnology Journal*, 15(11), 1374-1386. doi:10.1111/pbi.12722.

International Cotton Advisory Committee (ICAC) (2021). ICAC Data Portal. International Cotton Advisory Committee, Washington, DC.

Jia, X., Pang, C., Wei, H., et al. (2016). High-density linkage map construction and QTL analysis for earliness-related traits in *Gossypium hirsutum* L. *BMC Genomics*, 17, 909. https://doi.org/10.1186/s12864-016-3269-y.

Kuang, X. Q., and Yu, C. W. (2015). Generating cotton fiber length probability density function from fiber length parameters by finite mixture model. *The Journal of The Textile Institute*, 106(6), 655-662. doi:10.1080/00405000.2014.933516.

Kushanov, F. N., Turaev, O. S., Ernazarova, D. K., et al. (2021). Genetic diversity, QTL mapping, and marker-assisted selection technology in cotton (*Gossypium* spp.). *Frontiers in Plant Science*, 12, 779386. doi:10.3389/fpls.2021.779386.

Li, L., Zhao, S., Su, J., et al. (2017). High-density genetic linkage map construction by F_2 populations and QTL analysis of early-maturity traits in upland cotton (*Gossypium hirsutum* L.). *PLoS ONE*, 12(8), e0182918. https://doi.org/10.1371/journal.pone.0182918.

Littell, R. C., Milliken, G. A., Stroup, W. W., et al. (2006). SAS for Mixed Models (2nd ed.). SAS Institute, Cary, NC, USA.

Liu, D., Liu, F., Shan, X., et al. (2015). Construction of a high-density genetic map and lint percentage and cottonseed nutrient trait QTL identification in upland cotton (*Gossypium hirsutum* L.). *Molecular Genetics and Genomics*, 290(5), 1683-1700. doi:10.1007/s00438-015-1027-5.

Liu, G., Liu, J., Pei, W., et al. (2019). Analysis of the MIR160 gene family and the role of MIR160a_A05 in regulating fiber length in cotton. Planta, 250(6), 2147-2158. https://doi.org/10.1007/s00425-019 03271-7.

Manan, A., Zafar, M.M., Ren, M., et al. (2022). Genetic analysis of biochemical, fiber yield and quality traits of upland cotton under high-temperature. *Plant Production Science*, 25(1), 105-119. https://doi.org/10.1080/1343943X.2021.1972013.

McCall, L. L., Verhalen, L. M., and McNew, R. W. (1986). Multidirectional selection for fiber strength in upland cotton. *Crop Science*, 26, 744-750. doi:10.2135/cropsci1986.0011183X002600040024x.

Muthukumaran, S. (2016). Tree cotton (G. arboreum) in Babylonia. In E. Foietta, C. Ferrandi, E. Quirico, F. Giusto, M. Mortarini, J. Bruno, et al., (Eds.), Cultural & Material Contacts in the Ancient Near East: Proceedings of the International Workshop, December 1-2, 2014, Torino, Italy (pp. 98-105).

Nicholson, G. E. (1960). The production, history, uses and relationships of cotton (*Gossypium* spp.) in Ethiopia. *Economic Botany*, 14(1), 3-36. doi:10.1007/BF02859364.

Preetha, S., and Raveendren, T. S. (2008). Molecular marker technology in cotton. Biotechnology and Molecular Biology Reviews, 3(2), 32-45. doi:10.5897/BMBR2008.0004.

R Core Team (2014). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Sabev, P., Valkova, N., and Todorovska, E. G. (2020). Molecular markers and their application in cotton breeding: Progress and future perspectives. *Bulgarian Journal of Agricultural Science*, 26(4), 816-828.

Shang, L., Liu, F., Wang, Y., et al. (2015). Dynamic QTL mapping for plant height in Upland cotton (*Gossypium hirsutum*). *Plant Breeding*, 134(6), 703-712. doi:10.1111/pbr.12316.

Singh, R. K., and Chaudhary, B. D. (1985). Biometrical Methods in Quantitative Genetic Analysis. Kalyani Publishers, New Delhi, India. 318 pp.

Singh, B. (2001). Plant Breeding: Principles and Methods (6th ed.). Kalyani Publishers, New Delhi, India.

Smith, C. W., and Coyle, G. G. (1997). Association of fiber quality parameters and within-boll yield components in upland cotton. *Crop Science*, 37, 1775-1779.

Song, W., Wang, M., Su, W., et al. (2017). Genetic and phenotypic effects of chromosome segments introgressed from *Gossypium barbadense* into *Gossypium hirsutum*. *PLoS ONE*, 12(9), e0184882. https://doi.org/10.1371/journal.pone.0184882.

Song, C., Li, W., Pei, X., et al. (2019). Dissection of the genetic variation and candidate genes of lint percentage by a genome-wide association study in upland cotton. *Theoretical and Applied Genetics*, 132(7), 1991-2002. doi:10.1007/s00122-019-03333-0.

Stansfield, W. D. (1986). Theory and Problems of Genetics. McGraw-Hill Book Company, New York, USA.

Sun, F. D., Zhang, J. H., Wang, S. F., et al. (2012). QTL mapping for fiber quality traits across multiple generations and environments in upland cotton. *Molecular Breeding*, 30(1), 569 582. doi:10.1007/s11032-011-9645-z.

Sun, Z., Wang, X., Liu, Z., et al. (2017). Genome-wide association study discovered genetic variation and candidate genes of fiber quality traits in *Gossypium hirsutum* L. *Plant Biotechnology Journal*, 15(8), 982-996. doi:10.1111/pbi.12693.

Tan, Z., Zhang, Z., Sun, X., et al. (2018). Genetic map construction and fiber quality QTL mapping using the CottonSNP80K array in upland cotton. *Frontiers in Plant Science*, 9, 225. https://doi.org/10.3389/fpls.2018.00225.

Tyagi, A., Mir, Z. A., Almalki, M., et al. (2024). Genomics-assisted breeding: A powerful breeding approach for improving plant growth and stress resilience. *Agronomy*, 14(6), 1128. DOI: 10.3390/agronomy14061128.

Wang, B., Draye, X., Zhuang, Z., et al. (2017). QTL analysis of cotton fiber length in advanced backcross populations derived from a cross between *Gossypium hirsutum* and *G. mustelinum*. *Theoretical and Applied Genetics*, 130(6), 1297-1308. doi:10.1007/s00122-017-2889-1.

Wang, M., Tu, L., Lin, M., et al. (2017). Asymmetric subgenome selection and cis-regulatory divergence during cotton domestication. *Nature Genetics*, 49(4), 579-587. doi:10.1038/ng.3807.

Wang, B., Zhuang, Z., Draye, X., et al. (2020). Genome-wide quantitative trait loci reveal the genetic basis of cotton fibre quality and yield-related traits in a *Gossypium hirsutum* recombinant inbred line population. *Plant Biotechnology Journal*, 18(6), 1314-1327. https://doi.org/10.1111/pbi.13277.

Wang, F., Zhang, J., Chen, Y., et al. (2020). Identification of candidate genes for key fibre-related QTLs and derivation of favorable alleles in *Gossypium hirsutum* recombinant inbred lines with *G. barbadense* introgressions. *Plant Biotechnology Journal*, 18(3), 707-720. doi:10.1111/pbi.13237.

Xu, Y., and Crouch, J. H (2008). Marker-assisted selection in plant breeding: From publications to practice. *Crop Science*, 48(2), 391-407. doi:10.2135/cropsci2007.04.0191.

Yang, P., Sun, X., Liu, X., et al. (2022). Identification of candidate genes for lint percentage and fiber quality through QTL mapping and transcriptome analysis in an allotetraploid interspecific cotton CSSLs population. *Frontiers in Plant Science*, 13, 882051. https://doi.org/10.3389/fpls.2022.882051.

Yu, J. Z., Ulloa, M., Hoffman, S. M., et al. (2014). Mapping genomic loci for cotton plant architecture, yield components, and fiber properties in an interspecific (*Gossypium hirsutum* L. × *G. barbadense* L.) RIL population. *Molecular Genetics and Genomics*, 289(6), 1347-1367. doi:10.1007/s00438-014-0930-5.

Zhai, H., Gong, W., Tan, Y., et al. (2016). Identification of chromosome segment substitution lines of Gossypium barbadense introgressed in *G. hirsutum* and quantitative trait locus mapping for fiber quality and yield traits. *PLoS ONE*, 11(9), e0159101. doi: 10.1371/journal.pone.0159101.

751 Zhang, Y., Wang, L., Xin, H., et al. (2013). Construction of a high-density genetic map for sesame
752 based on large-scale marker development by specific length amplified fragment (SLAF)
753 sequencing. *BMC Plant Biology*, 13, 141. doi:10.1186/1471-2229-13-141.

754 Zhang, Z., Shang, H., Shi, Y., et al. (2016). Construction of a high-density genetic map by specific
755 locus amplified fragment sequencing (SLAF-seq) and its application to quantitative trait loci (QTL)
756 analysis for boll weight in upland cotton (*Gossypium hirsutum*). *BMC Plant Biology*, 16, 79.
757 doi:10.1186/s12870-016 0741-4.

758 Zhang, K., Kuraparthy, V., Fang, H., et al. (2019). High-density linkage map construction and QTL
759 analyses for fiber quality, yield and morphological traits using CottonSNP63K array in upland
760 cotton (*Gossypium hirsutum* L.). *BMC Genomics*, 20, 889. doi:10.1186/s12864-019-6214-z.

761 Zhang, Z., Zhang, T., Yang, J., et al. (2020). Genome-wide quantitative trait loci reveal the genetic
762 basis of cotton fibre quality and yield-related traits in a *Gossypium hirsutum* recombinant inbred
763 line population. *Plant Biotechnology Journal*, 18(6), 1314 1327. https://doi.org/10.1111/pbi.13277.

764 Zhang, C., Jiang, S., Tian, Y., et al. (2023). Smart breeding driven by advances in sequencing
765 technology. *Modern Agriculture*. DOI: 10.1002/moda.8.

766 Zhu, J., and Weir, B. S. (1998). Mixed model approaches for genetic analysis of quantitative traits.
767 Advanced Topics in Biomathematics, 1, 321-330.

768 Zumba, J., and Rodgers, J. (2016). Cotton micronaire measurements using small portable near
769 infrared (NIR) analyzers. Applied Spectroscopy, 70(5), 794-803. doi:10.1177/0003702816638227.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- AdditionalFigure1Aand1B.png
- AdditionalTable1ANOVA.xlsx
- AdditionalTable2.xlsx
- AdditionalTable3.xlsx
- AdditionalTable4.docx