QTL Mapping of Agronomic and Economic Traits for Four F2 Populations in Upland Cotton

Hongge Li
Chinese Academy of Agricultural Sciences Cotton Research Institute
Zhaoe Pan
Cotton Research Institute
Shoupu He
Cotton Research Institute
Yinhua Jia
Cotton Research Institute
Xiaoli Geng
Cotton Research Institute
Baojun Chen
Cotton Research Institute
Liru Wang
Cotton Research Institute
Baoyin Pang
Cotton Research Institute
Xiongming Du (dxm630723@163.com)
Chinese Academy of Agricultural Sciences Cotton Research Institute

Research

Keywords: upland cotton, yield traits, fiber quality traits, QTL mapping, simultaneous improvement

Posted Date: September 16th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-21705/v2

Version of Record: A version of this preprint was published on January 21st, 2021. See the published version at https://doi.org/10.1186/s42397-020-00076-y.

Abstract

Background: Upland cotton (*Gossypium hirsutum*) accounts for more than 90% of annual world cotton output due to its high yield potential. However, yield traits and fiber quality traits exhibit negative correlations in most cases. Here, we constructed four F_2 populations, using two normal lines and two introgression lines, for simultaneously detection the genetic basis underlying complex traits such as yield and fiber quality in upland cotton. Subsequently, the phenotyping of 8 agronomic and economic traits along with quantitative trait loci (QTL) mapping was implemented.

Results: Extensive phenotype variations and transgressive segregation were found across segregation populations. Four genetic maps were constructed with the length of 585.97cM, 752.45cM, 752.45cM and 1163.66cM. The mapping resulted in the identification 50 QTLs (27 were for fiber quality traits and 16 for yield traits) across four populations. Multiple QTLs having the common maker, such as qBW4 and qBW2, or residing in the same QTL cluster, such as qLP9 and qFL9-1, were prioritized for further research.

Conclusions: These findings will provide insight into the genetic basis of simultaneous improvement of yield and fiber quality in upland cotton breeding.

Introduction

Cotton represents the main source of natural textile fibers in the world and this is the most prevalent raw materials used in the textile industry (Wang et al. 2018). High yield and fine fiber-quality are prerequisites to meet the ever-increasing demand of the textile industry. Upland cotton (*Gossypium hirsutum*) accounts for more than 90% of global cotton production due to its high yield potential and broader adaptability but with moderate fiber quality, whereas *G. barbadense* produces exceptionally fine-quality fibers with lower fiber yield (Cai et al. 2014; Hu et al. 2019).

The majority of agronomic and economic traits, such as yield and fiber quality, are quantitative traits and controlled by multiple loci/genes. Moreover, environmental influence is substantial in the control and expression of these traits. Meanwhile, previous reports suggested a significantly negative correlations between fiber quality traits and yield traits (Wang et al. 2015; Liu et al. 2018; Zhang et al. 2019). Therefore, dissecting the genetic basis of yield and fiber quality is essential, and it would contribute to a simultaneous improvement for yield and fiber quality.

As a modern molecular genetic method, molecular markers have been widely applied in cotton in the last decade. Recently, the molecular markers get a great rapid development with the release of assembled genome sequences of *G. hirsutum* (Li et al. 2015; Zhang et al. 2015; Wang et al, 2018; Yang et al. 2019) and *G. barbadense* (Liu et al. 2015; Yuan et al. 2015). Numerous genetic linkage maps, including intraspecific map between *G. hirsutum* and interspecific map between *G.hirsutum* and *G.barbadense*, were constructed using restriction fragment length polymorphism (RFLP) , simple sequence repeats (SSR) and single nucleotide polymorphism (SNP), etc. According to CottonQTLdb (Release 2.3, Said et al. 2013, Said et al. 2015), thousands of quantitative trait loci (QTLs) for yield and fiber quality in cotton had

been detected. However, to date, the studies about simultaneous dissection of the genetic basis underlying complex traits and their genetic correlations in multiple upland cotton populations using QTL mapping remained few.

In the present study, four F_2 populations, which derived from the hybridization between two *G. hirsutum* normal lines (4133B and SGK9708) and two introgression lines (Suyuan04-3 and J02-247) were used. Subsequently, four corresponding genetic linkage maps were constructed using SSR markers. QTL mapping was implemented with the integration of genotypic data and phenotypic data of eight agronomic and economic traits, including yield and fiber quality. These findings will not only contribute to dissecting the genetic basis underlying yield and fiber quality and their genetic correlations, but also provide insight into the simultaneous improvement of yield and fiber quality in upland cotton breeding.

Materials And Methods

Plant Materials and Field Experiments

Two *G. hirsutum* normal lines (4133B and SGK9708), which are endowed with high yield potential but moderate fiber quality, and two introgression lines (Suyuan04-3 and J02-247), which are endowed with superior fiber quality, were as parents, respectively. SGK9708 was derived from CRI41, which is a widely planted cultivar with wide adaptability; 4133B was with greater combining ability and derived from the hybridization of SGK9708 and the offspring of Gan4104 and CZA(70)33; Suyuan04-3 was derived from the distant hybridization of [83-811×(86-1×G.*armourianum*)]; J02-247 was derived from the cross of Suyin45×Sukang310, with greater cotton boll as well as superior fiber length and strength. The cotton materials were provided by the National Mid-term Gene Bank for Cotton of China.

In 2014, the seeds of four parents were sown in Anyang, Henan province and four cross combinations, including 4133B×Suyuan04-3, 4133B×J02-247, SGK9708×J02-247 and SGK9708×4133B, were constructed. To facilitate the description below, four populations were termed as 4Su (4133B×Suyuan04-3), 4J (4133B×J02-247), SgJ (SGK9708×J02-247), and Sg4 (SGK9708×4133B). Winter breeding of cotton was carried out in Hainan province. In 2015, four F_2 populations (4Su, 4J, SgJ and Sg4), which consisting of 271, 248, 276 and 304 individuals respectively, were planted in Anyang. Each row was 7 m in length and 0.8 m apart, with 20-22 plants planted in each row. Field management practices were implemented according to local farming practices.

Trait Measurements and Statistical Analysis

At mid-September, all plants in four F_2 populations were used to investigate the plant height (PH). During the harvesting season, all the seed cotton was collected, and then boll weight (BW) and lint percentage (LP) were calculated after the seed cotton was weighed and ginned. Fiber quality traits, such as fiber length (FL), fiber strength (FS), fiber length uniformity (FU), micronaire (MIC) and fiber elongation (FE), were tested using an HVI1000 (Uster Technologies, Switzerland) in the Cotton Quality Supervision, Inspection and Testing Center, Ministry of Agriculture, Anyang, China,

The descriptive statistics, including the maximum value, minimum value, mean value, standard deviation, and coefficient of variation (CV), for the eight traits across four populations, were processed by Microsoft Excel 2013. The correlation matrix was calculated and visualized using corrplot package in R (Wei and Simko, 2016).

SSR Makers Analysis

Genomic DNA of individuals from F_2 populations and their parents was extracted from the young leaves tissue using a modified cetyltrimethyl ammonium bromide (CTAB) method (Paterson et al. 1993).

Polymorphism detection for four pairs of parents was run using 5713 SSR primers. The primers that amplify stable polymorphic products were selected for genotyping in F_2 population. The sequences of SSR primers were downloaded from CottonGen (https://www.cottongen.org/, Yu et al. 2014). In order to map SSRs to physical map, a local BLAST program was performed (Altschul et al. 1990). The sequences of SSRs were queried against the *G. hirsutum* genome sequences (Wang et al. 2018). The top 1 of blast-hit was selected for further analysis. The polymerase chain reaction (PCR), amplified products separating and silver staining were performed as detailed by Feng et al (2015).

Genetic map construction

The genetic linkage map was constructed using JoinMap 4.0 with a regression mapping method and logarithm of odds (LOD) threshold of 5.0. The Kosambi function was used to convert the recombination frequencies to map distances.

QTL mapping and analysis

Win QTL Cartographer 2.5 was applied to identify QTLs with the composite interval mapping (CIM) method. The main parameters were set as 1.0 cM for mapping step, 5 for control markers, and 1,000 for permutation tests. QTLs were declared significant if the corresponding LOD score was greater than 2.5. Meanwhile, the additive effect, dominant effect and R^2 (the percent of phenotypic variance explained by a QTL) were estimated. QTLs detected for the traits were named as follows: q-trait-linkage group No. (McCouch et al. 1988), where traits were PH, BW, LP, FL, FS, FU, FE and MIC. A graphic representation of the linkage groups and QTLs was created by MapChart 2.2 (Voorrips, 2002).

The action mode of QTL was represented by a dominance degree, i.e. an absolute value of dominant effect divided by additive effect (|D/A|; Stuber, 1987). It was additive if the dominance degree less than 0.2, partial dominance between 0.2 and 0.8, dominance between 0.81 and 1.2, over dominance more than 1.2.

The comparison between QTLs identified here and the CottonQTLdb database (Said et al. 2015) was carried out to determine whether QTLs were novel or detected before. Briefly, the QTLs in the present study that shared the same or overlapping confidence intervals with the QTLs in the database based on the common maker position were considered as QTLs identified in previous studies.

Results

Phenotypic variation of four F_2 populations

The phenotype of eight agronomic and economic traits across four F_2 populations was evaluated. As a result, the extensive phenotype variations and transgressive segregation were observed (Table 1 and Fig. 1), the transgressive segregation means that some individuals' phenotypic values were better than the superior parent and some' were worse than the inferior parent (Reyes 2019). The CV values indicated that there was a difference in variability between eight traits (Table 1). For PH, BW and LP, the CV value of LP was smaller (5.96%~7.98%), whereas the CV values of PH and BW were higher and similar (PH: 16.9%~21.95%; BW: 15.66%~19.7%). For FL, FS, FU, FE and MIC, the CV value of FU was minima (1.59%~2.61%) and MIC was maximum (13.87%~22%). Frequency distribution analysis depicted normal distribution of eight traits besides MIC (Fig. 1), suggesting that these traits were quantitative traits controlled by multiple genes and suitable for QTL mapping.

Correlation analysis

Correlation analysis between 32 sets of phenotypic data from eight traits across four populations illustrated the presence of significant correlations for different traits within and between populations (Fig. 2). BW showed significant negative correlation with LP ($-0.87 < r < -0.62$) within three populations (4Su, 4J, Sg4); while it depicted significant positive correlations with FS, FU, FE, MIC ($0.13<r<0.67$) within 4J and SgJ populations (Fig. 2). The negative correlations between BW and FL, FS, FU, FE ($-0.89<r<-0.82$) and the positive correlations between BW and MIC ($r=0.82$) were observed within 4Su populations (Fig. 2). In contrast, LP was positively correlated with FL, FS, FU, FE ($0.95<r<0.99$) and negatively correlated with MIC ($r=-0.90$) (Fig. 2). For fiber quality traits, the positive correlations between FL and FS, FU, FE; FS and FU, FE; FU and FE were observed within all four populations ($0.19 < r <0.998$) (Fig. 2).

PH and BW in 4J population were positively correlated with PH, LP, FL, FS, FU and FE in 4Su population ($0.25< r <0.88$), and negatively correlated with BW, MIC in 4Su population ($-0.776< r< -0.772$), respectively; as opposed to PH and BW, negative correlations ($-0.83<r < -0.15$) and positive correlations ($0.68<r<0.77$) between LP, FL in 4J population and corresponding traits in 4Su population were observed (Fig. 2). In addition, PH, LP, FE in Sg4 population was positively correlated with LP, FL, FS, FU, FE in 4Su population ($0.11< r <0.17$) and negatively correlated with BW ($-0.14<r< -0.13$), respectively; LP, FE, MIC in Sg4 population was positively correlated with PH, BW in 4J population ($0.14<r< 0.27$) and negatively correlated with LP, FL ($-0.24<r< -0.15$), respectively (Fig. 2).

Overall, within populations, the majority of correlations between two yield traits, BW and LP, were negative. In contrast, the majority of correlations among fiber qualities were positive, as well as between BW and fiber qualities (Fig. 2). The correlations between LP and fiber qualities were either positive or negative. The significant correlations between multiple traits among 4Su, 4J and Sg4 populations were observed (Fig. 2), suggesting the influence of common parent 4133B on traits.

Genetic map construction

5713 SSR primers were used to detect polymorphism for four pairs of parents, respectively. 739 polymorphism primers with clearly amplified bands were retained, including 203 polymorphism primers between 4133B and Suyuan04-3 (Additional file 5: Table S1a), 208 between 4133B and J02-407(Additional file 5: Table S1b), 158 between SGK9708 and J02-407 (Additional file 5: Table S1c), 170 between SGK9708 and 4133B (Additional file 5: Table S1d). The polymorphism rate of primers was 3.55%, 3.64%, 2.77% and 2.98% respectively.

Joinmap 4.0 software was employed to construct a genetic linkage map. For 4Su population, a total of 71 makers were assigned to 10 linkage groups (LGs) with a total map length of 585.97 cM (Table 2, Additional file 1: Fig. S1, Additional file 6: Table S2a). The average length of linkage groups was 58.6 cM, and the average distance of makers was 8.25 cM. The longest LG, LG9, contained the most makers (27), but half of LGs contained only three makers.

For 4J population, a map of 752.45 cM was constructed and 61 makers across 10 linkage groups were mapped (Table 2, Additional file 2: Fig. S2, Additional file 6: Table S2b). The average length of linkage groups was 75.2 cM, and the average distance of makers was 12.34 cM. LG7 contained most makers (21) and LG3 contained the least makers (3).

For SgJ population, 83 makers, approximately half of 158 polymorphism makers, were mapped in 15 linkage groups (Table 2, Additional file 3: Fig. S3, Additional file 6: Table S2c). The total length of the map was 855.04 cM and the average length of linkage groups was 57 cM. The highest adjacent maker interval was 21.46 cM on LG13 and least was 1.06 cM on LG14.

For Sg4 population, approximately one third of polymorphism makers (52/170) were assigned to 9 linkage groups, covering a genetic distance of 1163.66 cM (Table 2, Additional file 4: Fig. S4, Additional file 6: Table S2d). The average length of linkage groups was 129.3 cM, and the average distance of makers was 22.38 cM.

Mapping of QTLs

Win QTL Cartographer 2.5 was employed to identify the QTLs using the CIM algorithm for eight traits of four populations. As a result, a total of 50 QTLs were identified with 0.1%~59.24% R^2, of which 27 corresponds to fiber quality traits and 16 corresponds to yield traits. 23, 4, 8 and 15 QTLs were detected in 4Su, 4J, SgJ and Sg4 populations, respectively (Table 3, Fig 3). The LG9 in 4Su population harboured the most QTLs (13), following by LG6 (6) and LG1 (5) both in Sg4 population.

For PH, 7 QTLs identified, of which 6 in 4Su population, were all minor effect (0.11% < R^2< 4.02%; Table 3, Fig 3). The additive effect of two QTLs, qPH2-1 and qPH2-2, which with the higher R^2 (2.66% and 4.02%), were positive, indicating that the favourable alleles come from the parent Suyuan04-3. And the action mode of qPH2-1 and qPH2-2 were over dominance based on dominance degree value.

For BW, a total of 8 QTLs with 1.17%~9.31% R^2 were identified in 4J (1), SgJ (1) and Sg4 (6) (Table 3, Fig 3). It is noteworthy that both of the LGs harbouring one QTL in 4J (qBW4) and SgJ (qBW2) were anchored to A05 chromosome; meanwhile, a common SSR maker, NAU1255, was detected nearby the QTL interval. It was inferred that NAU1255 was a marker closely linked to BW. Furthermore, the directions of the additive effect and dominance effect were the same.

For LP, overall 8 QTLs with 1.68%~18.11% R^2 were identified in 4Su (2), 4J (2) and SgJ (4) (Table 3, Fig 3). The additive effect of two major QTLs, $qLP7_{4J}$ and qLP2, which with more than 10% R^2, were negative, indicating that the favourable alleles come from the parent J02-247. Moreover, the action mode of $qLP7_{4J}$ and qLP2 were over dominance and dominance, respectively.

For FL, the most QTLs (11) were detected. There were 6, 1 and 4 QTLs identified in 4Su, 4J and SgJ populations, respectively (Table 3, Fig 3). Multiple QTLs were in the same LG of a population, for example, qFL9-1, qFL9-2 and qFL9-3, which with 0.35% ~7.70% R^2, were in LG9 of 4Su population. Interestingly, both LG7 in 4Su population and LG6 in SgJ population were anchored to A13 chromosome. Meanwhile, the common SSR makers, BNL2449 and NAU1211, were detected nearby the interval of QTLs $qFL7_{4Su}$ and qFL6, hinting that BNL2449 and NAU1211 were closely linked concerning FL. In addition, the additive effect of QTLs qFL2-2 was positive, suggesting that the favourable alleles come from the male parent, Suyuan04-3 and J02-247, which is endowed with superior fiber quality.

For FS, a total of 5 QTLs were identified, 4 QTLs with R^2 of 2.95% ~7.15% in 4Su population and 1 major QTL with R^2 of 15.10% in Sg4 population (Table 3, Fig 3). The additive effect of 4 QTLs in 4Su population were positive, whereas 1 major QTL in Sg4 population was negative, implying that parent 4133B may not confer the favourable allele.

For FU, only two QTLs in the same LG of 4Su population with minor R^2 (0.10% ~1.21%) were identified (Table 3, Fig 3).

For FE, a total of 4 QTLs with 0.16% ~ 5.62% R^2 were detected in 4Su, SgJ and Sg4 populations (Table 3, Fig 3). The additive effect of one QTL, qFE8, was negative and action mode was additive, whereas, the other three QTLs were positive and over dominance.

For MIC, a total of 5 QTLs were detected across 3 LGs in 4Su and Sg4 populations (Table 3, Fig 3). As a major QTL, the R^2 of qMIC2, which in LG2 of Sg4 population was up to 59.24%, the other four QTLs R^2 were minor (0.15% ~6.29%). The dominance degree value of all QTLs but qMIC9-2 were up to 9.41~92.03, suggesting the action mode was over dominance.

There was a hotspot region in LG9 of 4Su population (Fig.3A). Three QTLs (qFL9-1, qFS9-1 and qFE9) were identified only at the position of 96.31cM; Further expansion of this region from 95.31cM to 105.81cM revealed presence of 8 QTLs corresponding 6 traits viz.PH (105.81cM), LP (95.31 cM), FL (96.31cM, 102.81 cM), FS (96.31cM, 101.81 cM), FE (96.31cM) and MIC(100.81cM). Therefore, this QTL

interval maybe an important genome region that affects agronomic and economic traits in cotton. At the same LG, two QTLs, qFU9-1 and qMIC9-1 were identified at the position of 41.71cM.

QTLs Comparison and Analysis

We compared the identified QTLs here and QTLs in CottonQTLdb database, the results showed that one-fifth of QTLs (10/50) overlapped with previously reported QTLs, illustrating the reliability of the QTL mapping in the present paper. Meanwhile, 40 novel QTLs were detected in our study. The overlapped 10 QTLs reportedly involved in FL (4), FS (2), PH (1), BW (1), LP (1) and FE (1) traits. There were the most identified QTLs both in the present research (11) and CottonQTLdb database (494) for FL, which perhaps will increase the probability of hit.

QTLs for different traits that shared the same or overlapping confidence intervals were considered to reside in QTL clusters. In the present study, a total of 9 QTL clusters were identified in 4Su (5), 4J (1) and Sg4 populations (3). The QTL cluster harbouring the most QTLs was above-mentioned hotspot region, with 8 QTLs for 6 traits, in LG9 of 4Su population (Fig.3A). There was another QTL cluster that harbouring QTLs for FU and MIC in the same LG (Fig.3A).

As we know, BW and LP represented yield traits, FL, FS, FU, FE and MIC represented fiber quality traits. With this prerequisite, the analysis of paired trait QTLs was employed. There were 19 paired trait QTLs within 6 paired traits (BW and FL, FE; LP and FL, FS, FU, FE) that exhibited significant medium or high positive correlations ($|r| > 0.3$) in the F_2 population. Among them, 6 paired trait QTLs had the same direction of addictive effect (Additional file 7: Table S3).

Discussion

To dissect the genetic basis underlying yield and fiber quality as well as their genetic correlations, two upland cotton normal lines (4133B and SGK9708) and two introgression lines (Suyuan04-3 and J02-247) were selected as parents respectively, and four populations were constructed. Among these populations, the female parents of 4Su, 4J and SgJ were high yield potential lines, and the male parents were superior fiber quality lines. Thus, the extensive phenotypic variation was observed in the cross combinations, whose parents are with distant kinship each other. Meanwhile, all traits exhibited normal distribution pattern across four F_2 populations (Table 1 and Fig. 1), suggesting that these traits were quantitative traits controlled by multiple genes

Furthermore, all traits exhibited transgressive segregation and many individuals with transgressive phenotype were found (Table 1 and Fig. 1). For example, all the median values of FL and FS in 4Su, 4J and SgJ populations were higher than or nearly 30, fiber reaching double-thirty quality values (FL $\geq$ 30 mm and FS $\geq$ 30 $cN \cdot tex^{-1}$) is generally considered as fine-quality. In plant breeding, transgressive segregation provides an adaptive advantage for traits (Reyes 2019). To a certain extent, high yield and fine-quality fibers are the outcome of adaptation for cotton. Therefore, it is not surprising that many instances of transgressive segregation were observed for these traits in F_2 populations. Furthermore,

some of these transgressive lines can be used to breed for high-quality fiber. At the same time, the above-mentioned phenomenon implied that the favourable alleles of fiber quality trait generally come from introgression lines' parents.

It is generally known that the quantitative traits are influenced by the environment. Therefore, to identify stable QTLs, the mapping populations are usually planted in multiple environments (Tang et al. 2015; Diouf et al. 2018; Zhang et al. 2019). However, multiple stable QTLs such as qBW4 and qBW2, were detected using four F_2 populations. Although these two QTLs were identified in 4J and SgJ populations, they had the common marker and their LGs anchored to the same chromosome. Thus these two QTLs could be considered as one QTL. In brief, this study provides an alternative method of detecting stable QTLs through multiple populations.

The phenomenon of QTLs cluster was consistent with previous studies, i.e. QTLs for fiber quality are clustered on the same chromosome; and the D09 chromosome, where the majority of makers in LG9 mapped, harbouring important loci regulating fiber quality traits (He et al. 2007; Qiao et al. 2019). These results illustrated that QTLs in clusters might be closely linked or have pleiotropic effects (Vikram et al. 2015; Zhao et al. 2016; You et al. 2019; Yuan et al. 2018), which explains the significant phenotypic correlations or linkage drag between related traits (Zhang et al. 2019). For paired trait QTLs, if they had the same QTL additive effect direction and showed significant medium or high positive correlations, it will be easy to simultaneously improve these traits (Zhang et al. 2019). In the present study, we identified 6 paired trait QTLs with significant positive correlation and additive mode of gene action (Additional file 7: Table S3). These results suggested valuable information for further simultaneous improvement of yield and fiber quality traits.

Based on the above conclusion, we found that qLP9 for LP and qFL9-1 as well as qFL9-2 for FL were in the same QTL cluster in LG9 of 4Su population. Furthermore, the high positive correlation and the same direction of the addictive effect between LP and FL were observed. Therefore, a further research plan is proposed using above mentioned QTLs cluster as a priority to include in a breeding program following fine mapping of QTL clusters via large scale segregating populations and gene-editing technology to break the negative correlation and further improve yield and fiber quality.

Conclusion

In this study, four F_2 populations were derived from the hybridization between two *G. hirsutum* normal lines and two introgression lines. Four corresponding genetic linkage maps were constructed. QTL mapping was implemented following the integration of phenotypic data of 8 agronomic and economic traits. A total of 50 QTLs across four populations, in which 27 were for fiber quality traits and 16 for yield traits, were detected. The QTLs in the same cluster, such as qLP9 for LP and qFL9-1 for FL, were prioritized for further research. These results will be helpful to dissect the genetic basis underlying yield and fiber quality, and lay a promising foundation for simultaneously improving of yield and fiber quality in upland cotton breeding programs.

Supplementary Information

Additional file 1: Fig.S1. The genetic map of 4Su population.

Additional file 2: Fig.S2. The genetic map of 4J population.

Additional file 3: Fig.S3. The genetic map of SgJ population.

Additional file 4: Fig.S4. The genetic map of Sg4 population.

Additional file 5: Table S1. The polymorphism primers information between four pairs of parents.

Additional file 6: Table S2. The linkage group of four populations.

Additional file 7: Table S3. Integrated analysis of paired trait QTLs and direction of additive.

Declarations

Acknowledgements

We would like to thank the anonymous reviewers for their valuable comments and helpful suggestions which help to improve the manuscript.

Authors' contributions

Li HG, Pan ZE, and Du XM designed the experiments and drafted the manuscript. Pan ZE and He SP carried out the molecular maker experiments. Jia YH and Geng XL participated in the design of the study and performed the statistical analysis. Li HG, Chen BJ, Wang LR, and Pang BY conducted the phenotypic evaluations and collected the data from the field. Li HG, Pan ZE, and He SP constructed the genetic maps and performed QTL mapping. All the authors read and approved the final manuscript.

Funding

This research was supported by the National Key R&D Program of China (2017YFD0101600).

Availability of data and materials

The data and materials for supporting the results of this article are included within the article and its supplementary material files.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

Altschul SF, Gish W, Miller W, et al. Basic local alignment search tool. *J Mol Biol.* 1990, 215(3):403-410. https://doi.org/10.1016/S0022-2836(05)80360-2

Cai C, Ye W, Zhang T, et al. Association analysis of fiber quality traits and exploration of elite alleles in Upland cotton cultivars/accessions (*Gossypium hirsutum* L.). *J Integr Plant Biol.* 2014, 56(1):51-62. https://doi.org/10.1111/jipb.12124

Diouf L, Magwanga RO, Gong W, et al. QTL Mapping of Fiber Quality and Yield-Related Traits in an Intra-Specific Upland Cotton Using Genotype by Sequencing (GBS). *Int J Mol Sci.* 2018, 19(2). https://doi.org/10.3390/ijms19020441

Feng H, Guo L, Wang G, et al. The Negative Correlation between Fiber Color and Quality Traits Revealed by QTL Analysis. *PLoS One.* 2015, 10(6). https://doi.org/10.1371/journal.pone.0129490

Gürbüz N, Eken NT, Ülger M, et al. Mapping of Quantitative Trait Loci for Antioxidant Molecules in Tomato Fruit: Carotenoids, Vitamins C and E, Glutathione and Phenolic Acids. *Plant Sci.* 2019:110393. https://doi.org/10.1016/j.plantsci.2019.110393

He D, Lin Z, Zhang X, et al. QTL mapping for economic traits based on a dense genetic map of cotton with PCR-based markers using the interspecific cross of *Gossypium hirsutum* x *Gossypium barbadense*. *Euphytica.* 2007, 153(1-2):181-197. https://doi.org/10.1007/s10681-006-9254-9

Hu Y, Chen J, Fang L, et al. *Gossypium barbadense* and *Gossypium hirsutum* genomes provide insights into the origin and evolution of allotetraploid cotton. *Nat Genet.* 2019, 51(4):739-748. https://doi.org/10.1038/s41588-019-0371-5

Li F, Fan G, Lu C, et al. Genome sequence of cultivated Upland cotton (*Gossypium hirsutum* TM-1) provides insights into genome evolution. *Nat Biotechnol.* 2015, 33(5):524-530. https://doi.org/10.1038/nbt.3208

Liu R, Gong J, Xiao X, et al. GWAS Analysis and QTL Identification of Fiber Quality Traits and Yield Components in Upland Cotton Using Enriched High-Density SNP Markers. *Front Plant Sci.* 2018, 9. https://doi.org/10.3389/fpls.2018.01067

Liu X, Zhao B, Zheng H, et al. *Gossypium barbadense* genome sequence provides insight into the evolution of extra-long staple fiber and specialized metabolites. *Sci Rep.* 2015,

5https://doi.org/10.1038/srep14139

Ma Z, He S, Wang X, et al. Resequencing a core collection of upland cotton identifies genomic variation and loci influencing fiber quality and yield. *Nat Genet.* 2018, 50(6):803-813. https://doi.org/10.1038/s41588-018-0119-7

McCouch SR, Kochert G, Yu ZH, et al. Molecular mapping of rice chromosomes. *Theor Appl Genet.* 1988, 76(6):815-829. https://doi.org/10.1007/bf00273666

Paterson AH, Brubaker CL, Wendel JF. A rapid method for extraction of cotton (*Gossypium* spp.) genomic DNA suitable for RFLP or PCR analysis. *Plant Mol Biol Rep.* 1993, 11(2):122-127. https://doi.org/10.1007/bf02670470

Qiao W, Yan G, Shi J, et al. QTL Mapping of Fiber Quality in Low-Generation Upland Cotton Populations and Functional Annotation of Candidate Genes. *Cotton Science.* 2019, 31(4):282-296. https://doi.org/10.11963/1002-7807.qwqhq.20190530

Reyes BGL. Genomic and epigenomic bases of transgressive segregation - New breeding it paradigm for novel plant phenotypes. *Plant Sci.* 2019, 288:10. https://doi.org/10.1016/j.plantsci.2019.110213

Said JI, Knapka JA, Song M, et al. Cotton QTLdb: a cotton QTL database for QTL analysis, visualization, and comparison between *Gossypium hirsutum* and *G.hirsutum* x *G. barbadense* populations. *Mol Genet Genomics.* 2015, 290(4):1615-1625. https://doi.org/10.1007/s00438-015-1021-y

Stuber CW, Edwards MD, Wendel JF. Molecular marker-facilitated investigations of quantitative trait loci in maize. II. Factors influencing yield and its component traits. *Crop Sci.* 1987, 27(4):639-648. https://doi.org/10.2135/cropsci1987.0011183X002700040006x

Tang S, Teng Z, Zhai T, et al. Construction of genetic map and QTL analysis of fiber quality traits for Upland cotton (*Gossypium hirsutum* L.). *Euphytica.* 2015, 201(2):195-213. https://doi.org/10.1007/s10681-014-1189-y

Voorrips RE. MapChart: Software for the graphical presentation of linkage maps and QTLs. *J Hered.* 2002, 93(1):77-78. https://doi.org/10.1093/jhered/93.1.77

Wang HT, Huang C, Guo HL, et al. QTL Mapping for Fiber and Yield Traits in Upland Cotton under Multiple Environments. *PLoS One.* 2015, 10(6):14. https://doi.org/10.1371/journal.pone.0130742

Wang M, Tu L, Yuan D, et al. Reference genome sequences of two cultivated allotetraploid cottons, *Gossypium hirsutum* and *Gossypium barbadense*. *Nat Genet.* 2018. https://doi.org/10.1038/s41588-018-0282-x

Yang Z, Ge X, Yang Z, et al. Extensive intraspecific gene order and gene structural variations in upland cotton cultivars. *Nat Commun.* 2019, 10(1):2989. https://doi.org/10.1038/s41467-019-10820-x

You W, Feng Y, Shen Y, et al. Silencing of NbCMT3s has Pleiotropic Effects on Development by Interfering with Autophagy-Related Genes in *Nicotiana benthamiana*. *Plant Cell Physiol.* 2019, 60(5):1120-1135. https://doi.org/10.1093/pcp/pcz034

Yu J, Jung S, Cheng C-H, et al. CottonGen: a genomics, genetics and breeding database for cotton research. *Nucleic Acids Res.* 2014, 42(D1):D1229-D1236. https://doi.org/10.1093/nar/gkt1064

Yuan D, Tang Z, Wang M, et al. The genome sequence of Sea-Island cotton (*Gossypium barbadense*) provides insights into the allopolyploidization and development of superior spinnable fibres. *Sci Rep.* 2015, 5. https://doi.org/10.1038/srep17662

Yuan X, Huang P, Wang R, et al. A Zinc Finger Transcriptional Repressor Confers Pleiotropic Effects on Rice Growth and Drought Tolerance by Down-Regulating Stress-Responsive Genes. *Plant Cell Physiol.* 2018, 59(10):2129-2142. https://doi.org/10.1093/pcp/pcy133

Zhang T, Hu Y, Jiang W, et al. Sequencing of allotetraploid cotton (*Gossypium hirsutum* L. acc. TM-1) provides a resource for fiber improvement. *Nat Biotechnol.* 2015, 33(5):531-537. https://doi.org/10.1038/nbt.3207

Zhang Z, Li J, Jamshed M, et al. 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 Biotechnol J.* 2020, 18(1):239-253. https://doi.org/10.1111/pbi.13191

Zhao X, Daygon VD, McNally KL, et al. Identification of stable QTLs causing chalk in rice grains in nine environments. *Theor Appl Genet.* 2016, 129(1):141-153. https://doi.org/10.1007/s00122-015-2616-8

Tables

Tables can be found in the Supplemental files section.

Figures

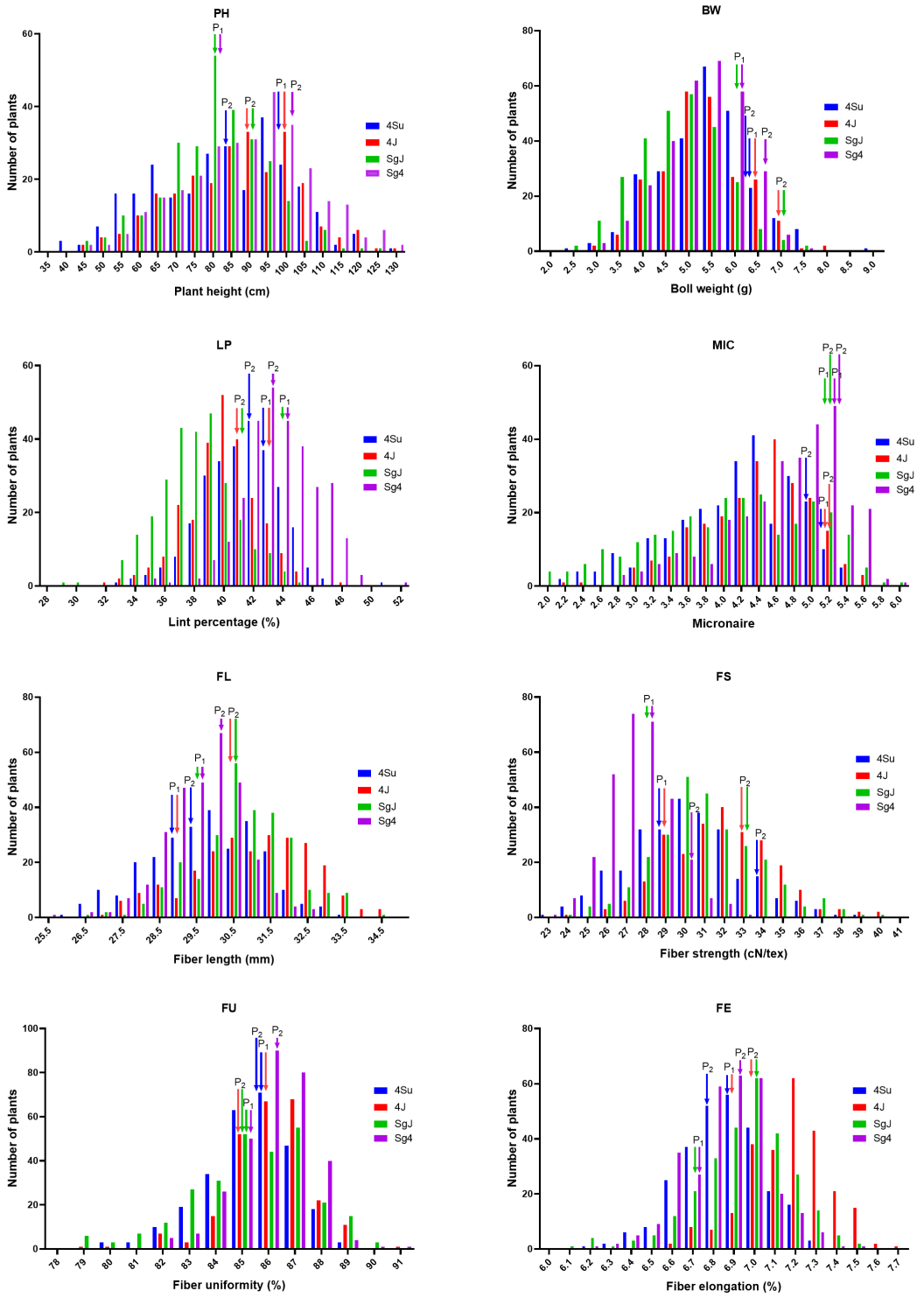

Figure 1

Histogram of the frequency distribution for 8 traits across 4 populations. plant height (PH), boll weight (BW), lint percentage (LP), fiber length (FL), fiber strength (FS), fiber length uniformity (FU), micronaire (MIC), fiber elongation (FE).

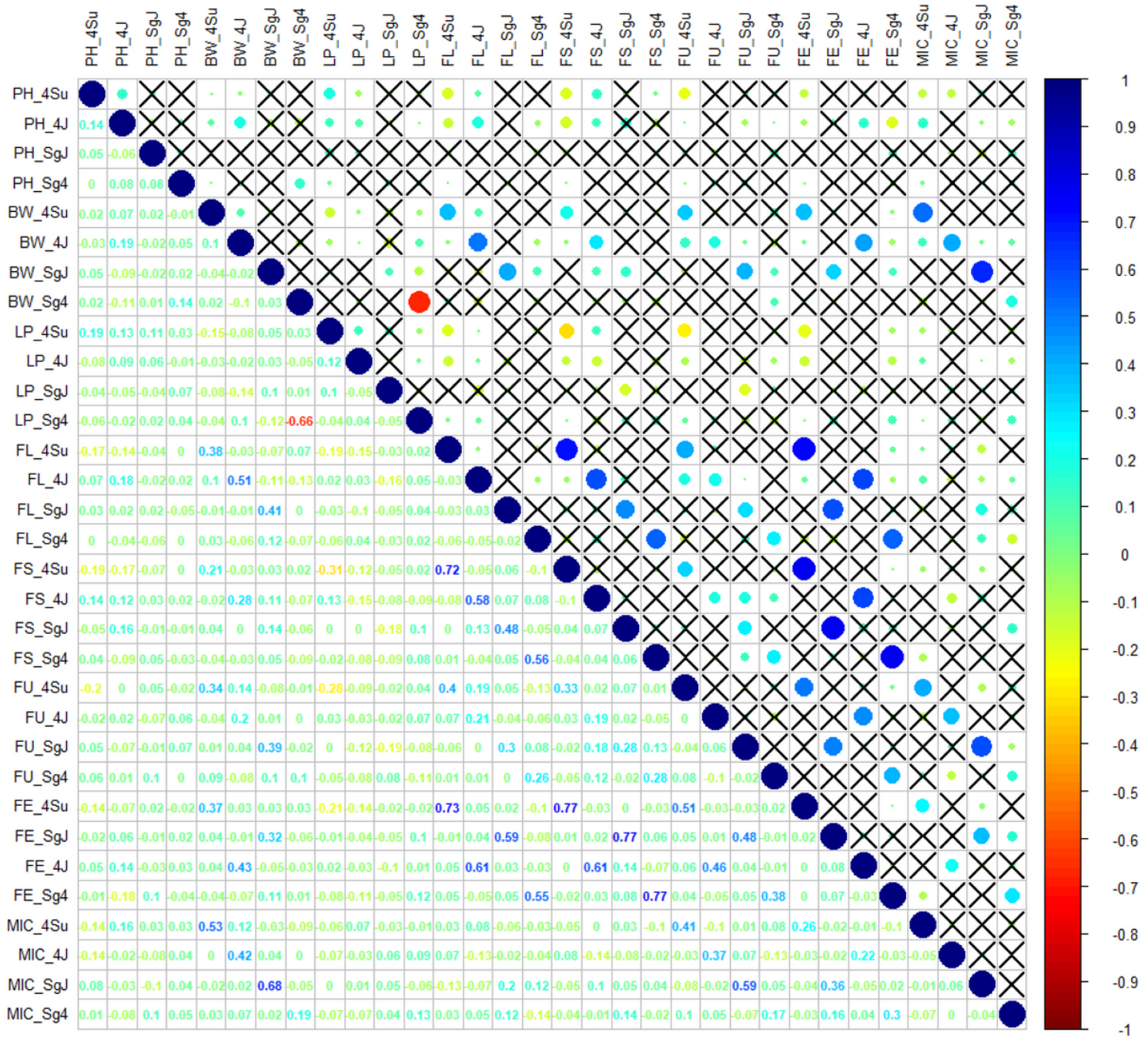

Figure 2

Correlation analysis among 8 traits. The magnitude of the correlation is indicated by different colors of number at the left diagonal and circles at the right diagonal. The circles marked by cross indicate no significant correction was observed between traits (P > 0.05).

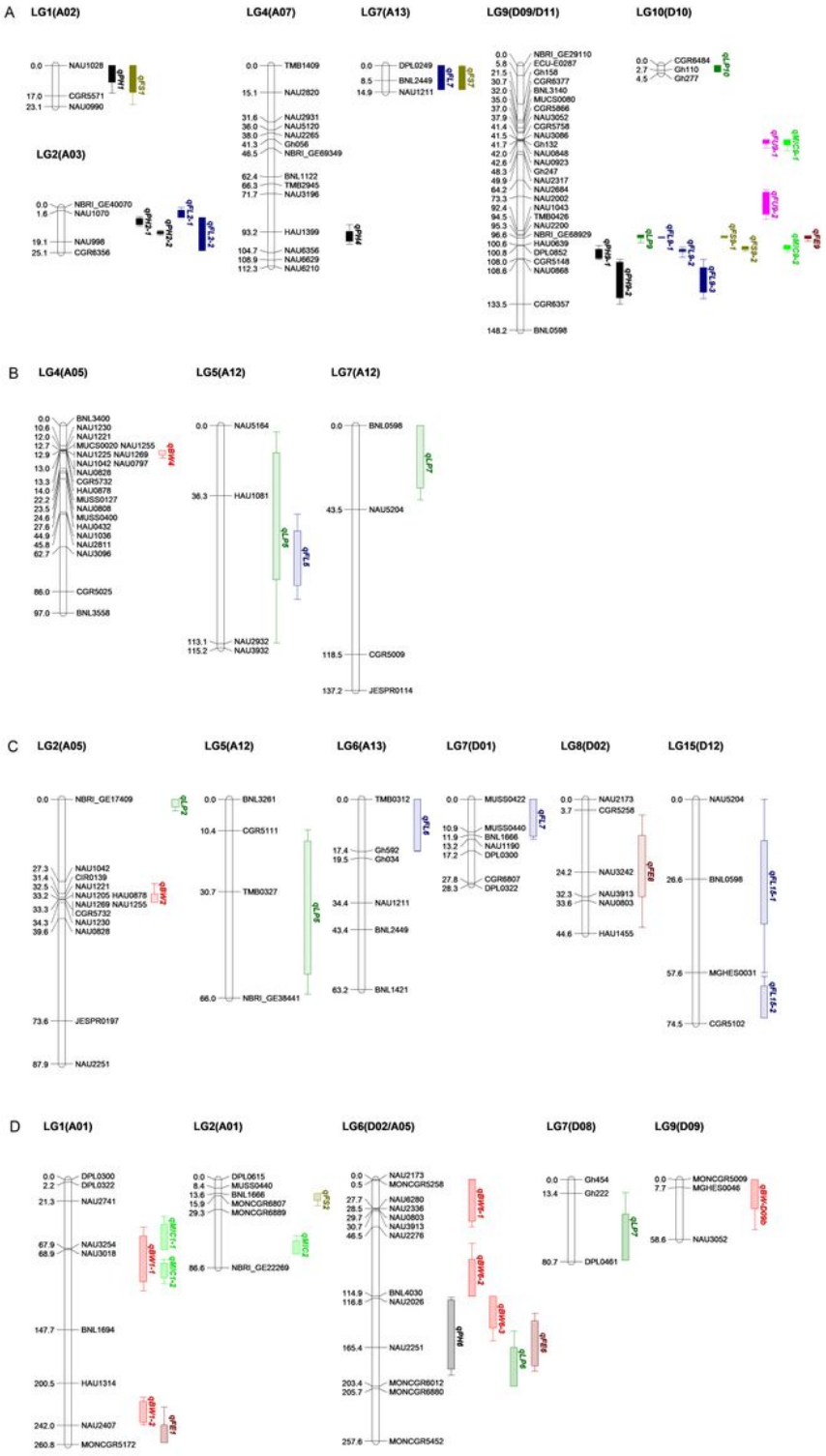

Figure 3

QTLs identified for 8 traits across 4 populations. A: 4Su population; B: 4J population; C: SgJ population; D: Sg4 population.

Supplementary Files

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

- Table1.xlsx
- Table2.xlsx
- Table3.xlsx
- TableS1markerinfo.xlsx
- TableS2allLGs.xlsx
- TableS3pairedtraits.xlsx
- Fig.S14133su.tif
- Fig.S24133j.tif
- Fig.S3sgj.tif
- Fig.S4sg4133.tif