Construction of a High-Density Genetic Map via Genome Resequencing and QTL Detection for Key Fiber Traits in Asiatic Cotton

Yaohua Li  
Southwest University  
https://orcid.org/0000-0003-1385-1229

Tong Mo  
Southwest University

Lingfang Ran  
Southwest University

Jianyan Zeng  
Southwest University

Chuannan Wang  
Southwest University

Ziman Zhong  
Southwest University

Aimin Liang  
Southwest University

Yuehua Xiao  
xiaoyuehua@swu.edu.cn  
Southwest University

Research article

Keywords: Asiatic cotton (Gossypium arboretum), Resequencing, Genetic map, Fiber quality traits, Fiber yield traits, QTL

DOI: https://doi.org/10.21203/rs.3.rs-53461/v1

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

**Background:** Asiatic cotton (*Gossypium arboreum*, genome A$_2$) is one of diploid cotton species producing spinnable fibers. However, few studies on the genetic mechanism of key fiber traits of Asiatic cotton have been reported. Sequencing technology advancement and the release of Asiatic cotton genome made it possible to construct a high-density SNP genetic map and further untapped QTL detection.

**Results:** The Asiatic cotton cultivars SXY No.1 and CSLZ were crossed to develop a recombinant inbred line (RIL) population with 189 lines. Whole genome resequencing technology was employed to construct a high-density genetic map that covered 1980.17 cM with an average distance of 0.61 cM between adjacent markers. Based on fiber quality and yield component trait data from three environments, a total of 177 QTL were identified for 8 key fiber traits explaining 5.0-37.4% of the phenotypic variance. Besides, 48 stable QTL, including 15 for upper quartile length (UQL), 18 for fiber fineness (FF), 1 for immature fiber content (IFC), 4 for fiber neps count (FNC), 3 for lint percentage (LP), 7 for seed index (SI), were detected in more than one environment.

**Conclusions:** Using a RIL population and whole genome resequencing strategy, this study presented a high-density genetic map of *G. arboreum* and identified 48 stable QTL for 6 key fiber traits (UQL, FF, IFC, FNC, LP, SI). Our work laid solid foundation for subsequent fine mapping of QTL for key fiber traits and cloning of controlling genes.

**Background**

Cotton is the most important fiber crop in the world. Cultivated species of genus *Gossypium* include A-genome diploids *G. herbaceum* (A$_1$) and *G. arboreum* (A$_2$) and AD-genome allotetraploids *G. hirsutum* (AD$_1$) and *G. barbadense* (AD$_2$) which produce spinnable fiber [1]. It was believed that the allotetraploids arose from an interspecific hybridization between an A-genome diploid and D-genome taxon within the last 1–2 million years ago [1, 2]. Presently, *G. hirsutum* is most widely planted with more than 95% worldwide production due to its superior fiber quality and high yield [3]. However, the diversity of *G. hirsutum* germplasm base is relatively narrow [4]. In contrast, the cultivation scale of diploid cotton is very limited, mainly in India, Pakistan and China [5], but they are extensively deployed by many cotton breeders for a series of genetic and biotechnological studies owing to diverse agronomic and fiber properties.

Asiatic cotton (*G. arboreum* L.) derived in the Indian subcontinent and has been cultivated for 5000 years [6]. During long-term natural and artificial selection, Asiatic cotton possesses many favorable traits that are not found in the allotetraploids cultivars. For example, adaptive characteristics like remarkable drought tolerance [7], enhanced resistance against insect pests like thrips [8], bollworms and aphids and diseases such as black arm, root rot and leaves reddening [9], complete immunity to cotton leaf curl virus [10]. Some of *G. arboreum* cultivars produce fibers with high strength and seeds with high oil content and seed index [9]. Fiber related properties are generally polygenic traits. Compared to allotetraploids, diploid cotton may facilitate to study these characters for reducing gene duplication and genetic redundancy [11]. Future efforts could utilize favorable traits found in Asiatic cotton species as donor in introgressive improvement of allotetraploid cotton. And molecular markers are expected to accelerate breeding progress with which traits can be introgressed [12].

In recent years, SNP (single nucleotide polymorphism) molecular marker has become popular in high-density genetic map construction, QTL mapping and marker assisted selection (MAS) due to its abundance in plant genome and advancement in next generation sequencing (NGS) [13–15]. And many NGS-related approaches have been developed to identify SNPs in cotton, such as whole genome resequencing [16, 17], restriction site-associated DNA sequencing (RAD-seq) [18, 19], specific locus amplified fragment sequencing (SLAF-seq) [20–22] and genotyping-by-sequencing (GBS) [23, 24]. In addition, the release of genome sequences of cotton species including *G. arboreum* [16], *G. raimondii* [25], *G. hirsutum* [26] and *G. barbadense* [27] made it easy to construct genetic map and detect QTL via NGS technology.

Recently, considerable work of high-density genetic map construction was performed in *G. hirsutum* and *G. barbadense* with resequencing, RAD-seq, SLAF-seq or GBS, and many fiber quality and lint yield trait QTL have been identified [17, 18, 20, 24]. Consequently, it is effective for identifying untapped alleles of cotton quantitative traits to use a strategy of combining NGS and...
ILs. For diploid cottons, these researches in high-density genetic map and QTL mapping relatively delayed. Du et al. [16] resequenced 243 lines of G. arboretum and G. herbaceum and identified 98 significant peak associations for 11 agronomically traits combined GWAS analysis, but candidate genes for fiber quality traits were not reported. Before this study, researchers were all employed traditional molecular markers, such as RFLP, AFLP, TRAP and SSR, to construct interspecific and intraspecific genetic map of Asiatic cotton and to map QTL [28–30]. These genetic maps cannot meet the demands of accurate QTL mapping due to their large distances between adjacent DNA markers. Therefore, a high-density genetic map within G. arboreum and corresponding QTL detection are currently in demand.

In this study, we developed an Asiatic cotton recombinant inbred line (RIL) population of 189 F2:6 lines from a cross SXY No.1 · CSLZ. Whole genome resequencing was applied to genotype RILs, which was then used to construct the high-density genetic map and detect QTL of key fiber traits. Our results will facilitate to fine map the QTLs for key fiber traits in Asiatic cotton and subsequently to clone the controlling genes.

Results

Fiber trait evaluation

Because there were missing data for some lines across three environments, we finally used trait data of 179 lines for phenotypic analysis and subsequent QTL mapping. Descriptive statistics for all key fiber traits i.e. upper quartile length (UQL), fiber fineness (FF), maturity ratio (MR), immature fiber content (IFC), fiber neps count (FNC), fiber neps mean size (FNMS), lint percentage (LP) and seed index (SI) across three environments were summarized in Additional file 9: figure S1 and Additional file 1: Table S1. The absolute values of kurtosis and skewness for most traits were <1.0 in three environments, except for the kurtosis values for UQL-2018CQ, FNC-2019CQ, LP-2018CQ and LP-2018HN, suggesting that most traits were normally distributed. Besides, all eight traits of population showed transgressive segregation.

Correlation analysis was conducted among the eight traits (Additional file 2: Table S2). Most traits showed significant correlations with other traits except for UQL-FNMS, FF-LP, FNMS-LP. One-way ANOVA showed that most traits had significant genetic and environmental effects ($p < 0.01$) except for genetic effect of FNMS (Additional file 3: Table S3).

Whole-genome sequencing and SNP identification

Sequencing data were generated from two parents and their RIL population. In total, 3808.40 million clean reads were obtained including 255.62 million clean reads for parent SXY No.1, 233.81 million clean reads for parent CSLZ, and 3318.97 million clean reads for their offspring respectively. The average sequencing depth of SNP markers were 21-fold for SXY No.1, 19-fold for CSLZ and 2-fold for their progeny. Among these reads data, the percentages of reads properly mapped on reference genome for SXY No.1, CSLZ and RILs were 98.48, 98.20 and 98.16% respectively (Additional file 4: Table S4). A total of 594, 416 SNPs were detected between two parents in our study. According to the character of RILs, 130,696 (21.99% of) SNPs with genotype aa ´ bb was used for further analysis and 62,593 SNPs were retained after subsequent multiple filtering. After final segregation filtration of bins, 45,468 SNPs which were divided into 3286 recombinant bins were employed to construct the high-density genetic map (Additional file 5: Table S5).

Genetic map construction

By genetic linkage analysis, a high-density genetic map containing 3286 bin markers was constructed, which covered 1980.17 cM with an average distance of 0.61 cM between adjacent bin markers (Fig. 1; Table 1; Additional file 10: Supplement S1). 13 linkage groups were corresponding to 13 chromosomes respectively in this genetic map. The largest chromosome was Chr07, consisting of 392 bins covering 207.11 cM, with an average bin interval was 0.53 cM. The shortest chromosome was Chr02, containing 135 markers spanning 106.5 cM, with an average bin interval of 0.79 cM. Chr09 harbored the largest gap that was 12.62 cM. Besides, 99.41% of the intervals between adjacent bins were less than 5 cM, indicating that bin markers were well-distributed on the genome (Fig. 1; Table 1).

Segregation distortion analysis
Among 3286 mapped loci, 446 (13.57%) showed segregation distortion \((p < 0.05)\) (Table 1). The segregation distortion markers (SDMs) were unevenly distributed over the genome, forming 26 segregation distorted regions (SDRs) (Fig. 1). Chr09 had the largest number of distorted loci (97), accounting for 37.75% of the total markers of Chr09. While Chr04 contained only 4 SDMs with the smallest proportion of 1.89%. A total of 323 SDMs (72.42%) favored SXY No.1 alleles while 123 SDMs (27.58%) favored CSLZ.

**Table 1** Characteristics of the genetic map

| Linkage Groups | Total Bin Marker | Total SNP Marker | Total Distance (cM) | Average Distance (cM) | Max Gap (cM) | Gap < 5 cM (%) | SDM | SDM rate (%) | SDR |
|----------------|-----------------|-----------------|--------------------|-----------------------|-------------|----------------|-----|--------------|-----|
| LG01           | 227             | 3386            | 128.95             | 0.57                  | 6.97        | 99.12          | 50  | 22.03        | 3   |
| LG02           | 135             | 1324            | 106.5              | 0.79                  | 5.28        | 99.25          | 18  | 13.33        | 2   |
| LG03           | 285             | 3574            | 115.55             | 0.41                  | 4.32        | 100.00         | 24  | 8.42         | 2   |
| LG04           | 212             | 2277            | 136.67             | 0.64                  | 6.27        | 99.53          | 4   | 1.89         | 0   |
| LG05           | 323             | 3622            | 202.93             | 0.63                  | 6.47        | 99.69          | 15  | 4.64         | 2   |
| LG06           | 249             | 2736            | 173.74             | 0.70                  | 10          | 99.60          | 67  | 26.91        | 3   |
| LG07           | 392             | 6661            | 207.11             | 0.53                  | 8.41        | 99.74          | 11  | 2.81         | 1   |
| LG08           | 217             | 2892            | 141.17             | 0.65                  | 5.83        | 99.54          | 5   | 2.30         | 0   |
| LG09           | 244             | 3388            | 176.9              | 0.72                  | 12.62       | 98.77          | 97  | 39.75        | 4   |
| LG10           | 331             | 5262            | 156.38             | 0.47                  | 3.93        | 100.00         | 63  | 19.03        | 2   |
| LG11           | 274             | 4720            | 155.21             | 0.57                  | 6.62        | 99.27          | 31  | 11.31        | 4   |
| LG12           | 244             | 3728            | 151.32             | 0.62                  | 7.69        | 99.18          | 12  | 4.92         | 1   |
| LG13           | 153             | 1898            | 127.74             | 0.83                  | 7.94        | 98.68          | 49  | 32.03        | 2   |
| Total          | 3286            | 45468           | 1980.17            | 0.61                  | 12.62       | 99.41          | 446 | 13.57        | 26  |

LG linkage group, SDM segregation distorted marker, SDR segregation distorted region

**Collinearity between the genetic and the physical map**

In order to assess the quality of the genetic map, collinearity between genetic and physical map was conducted (Figure 2). Most genetic loci were in accordance with their positions on the reference genome sequence of *G. arboreum* except that 5 loci of LG08 corresponded to the 5 loci on Chr02 and 60 loci located on chr1g which were not mapped to chromosomes (Additional file 6: Table S6). Furthermore, 1980.17 cM corresponded to 1.46 GB, which covered 99.24% of the genome and all chromosomes showed more than 95% coverage (Additional file 7: Table S7).

**QTL mapping**

A total of 177 QTL for key fiber traits were identified in this study (Additional file 8: Table S8). 136 QTL for 6 fiber quality traits (UQL, FF, MR, IFC, FNC and FNMS) were detected. Phenotypic variance explained (PVE) by these QTL was 5.0-17.7%, and LOD scores ranged from 2.01-7.59. 72 QTL which had positive additive effects derived from SXY No.1, while 64 QTL which had negative additive effects derived from CSLZ.

41 QTL for 2 fiber yield traits (LP, SI) were detected. Their LOD scores ranged from 2.04-18.2, and they explained 5.1-37.4% phenotypic variance. SXY No.1 contributed 30 favorable alleles while CSLZ contributed 11. Besides, 48 QTL were detected with
stability more than one environment including 38 fiber quality related QTL and 10 fiber yield related QTL (Table 2).

**Upper quartile length QTL**

Thirty-seven upper quartile length QTL were found on 12 chromosomes, with LOD scores ranging from 2.02 to 6.32 and explaining 5.1-15% of the phenotypic variance (Additional file 8: Table S8). The favorable alleles of 25 QTL came from SXY No.1, and 12 came from CSLZ. Fifteen QTL were identified in two or more environments, but only two (qUQL04.3, qUQL11.3) were identified in all three (Table 2; Fig. 1). Six QTL (qUQL01.4, qUQL03.1, qUQL05.2, qUQL06.3, qUQL11.1 and qUQL11.3) explained more than 10% of the phenotypic variance.

**Fiber fineness QTL**

Thirty-two fiber fineness QTL were also identified on 12 chromosomes, with LOD scores ranging from 2.03 to 7.59, and PVE values ranging from 5.1 to 17.7% (Additional file 8: Table S8). Among these QTL, the favorable alleles derived from two parents were all 16. Eighteen QTL were identified in two or more environments and eight QTL (qFF02.1, qFF02.4, qFF05.1, qFF05.2, qFF05.3, qFF10.2, qFF11.1, qFF13.2) were detected in three environments (Table 2; Fig. 1). Ten QTL (qFF02.4, qFF02.5, qFF03.4, qFF05.3, qFF05.4, qFF06.1, qFF07.1, qFF10.2, qFF11.2 and qFF13.2) explained more than 10% of the phenotypic variance.

**Maturity ratio QTL**

Eighteen maturity ratio QTL were found on 12 chromosomes, with LOD scores ranging from 2.15 to 5.6, and explaining 5.4-13.4% of the phenotypic variance (Additional file 8: Table S8). SXY No.1 contributed 10 favorable alleles for maturity ratio while CSLZ contributed 8. Four QTL (qMR02.1, qMR 04.1, qMR07.2, qMR13.1) explained more than 10% of the phenotypic variance.

**Immature fiber content QTL**

Twenty immature fiber content QTL were found on 12 chromosomes, with LOD scores ranging from 2.01 to 5.86, and explaining 5-14% of the phenotypic variance (Additional file 8: Table S8). The favorable alleles of 9 QTL originated from SXY No.1, and 11 originated from CSLZ. However, only one QTL (qIFC13.1) with stability was detected in two environments (Table 2; Fig. 1). Two QTL (qIFC04.3 and qIFC13.1) explained more than 10% of the phenotypic variance.

**Fiber neps count QTL**

Sixteen fiber neps count QTL were detected on 10 chromosomes, with LOD scores ranging from 2.09 to 5.87, and explaining 5.2-14% of the phenotypic variance (Additional file 8: Table S8). Among them, favorable alleles for 6 QTL were contributed by SXY No.1, and the rest were from CSLZ. Four QTL (qFNC01.1, qFNC07.1, qFNC07.5, qFNC11.1) were identified in two or more environments but only one QTL (qFNC07.1) was detected in all three, explaining 14% of the phenotypic variance (Table 2; Fig. 1).

**Fiber neps mean size QTL**

Thirteen fiber neps mean size QTL were identified on 9 chromosomes, with LOD scores ranging from 2.03 to 5.35, and explaining 5.1-12.9% of the phenotypic variance (Additional file 8: Table S8). Favorable alleles of 6 QTL were conferred by SXY No.1. and the others by CSLZ. Only one QTL qFNMS04.1 explained more than 10% of the phenotypic variance.

**Lint percentage QTL**

Fourteen lint percentage QTL were identified on 8 chromosomes, with LOD scores ranging from 2.08 to 18.2, and explaining 5.2-37.4% of the phenotypic variance (Additional file 8: Table S8). SXY No.1 contributed 8 favorable alleles for lint percentage whereas CSLZ contributed 6. Three QTL (qLP11.1, qLP12.1, qLP13.1) were identified in two or more environments. However, only QTL qLP13.1 was identified in three environments, explaining 37.4% of the phenotypic variance (Table 2; Fig. 1).

**Seed index QTL**
Twenty-seven seed index QTL were detected on 11 chromosomes, with LOD scores ranging from 2.04 to 7.59, and explaining 5.1-17.7% of the phenotypic variance (Additional file 8: Table S8). The favorable alleles of 22 QTL derived from SXY No.1, and the rest 5 QTL were from CSLZ. Seven QTL (qSI01.3, qSI04.1, qSI07.1, qSI07.2, qSI07.3, qSI08.3, qSI09.4) were found in two or more environments whereas only QTL qSI04.1 was identified in three environments, explaining 5.7% of the phenotypic variance (Table 2; Fig. 1). Six QTL (qSI01.3, qSI07.2, qSI07.3, qSI08.3, qSI09.4, qSI12.1) explained more than 10% of the phenotypic variance.

QTL clusters and hotpots

QTL clusters and hotpots were defined as regions contained multiple QTL within approximately 20 cM, for different and same traits, respectively [24, 31]. In this study, a total of 47 QTL clusters comprising 131 QTL were identified across 13 chromosomes whereas no hotspot was detected (Table 3). Chromosome 7 had the highest number of clusters (8 clusters for Chr07-cluster-1, Chr07-cluster-2, Chr07-cluster-3, Chr07-cluster-4, Chr07-cluster-5, Chr07-cluster-6, Chr07-cluster-7, Chr07-cluster-8). In addition, every QTL cluster included at least two QTL for different traits and Chr12-cluster-1 possessed the highest number of QTL (6 QTL for qFF12.1, qMR12.1, qIFC12.1, qLP12.1, qUQL12.2, qSI12.1). Notably, twenty-six clusters contained stable QTL which could be key regions of fiber traits regulatory genes.

Discussion

High-density genetic map construction

A high-density genetic map is very important for the detection of new QTL, gene mapping and maker assisted selection in cotton. In previous studies, traditional molecular markers, such as RFLP, AFLP, TRAP and SSR, were widely used in most interspecific and intraspecific genetic map construction of Asiatic cotton [28-30]. However, most of these genetic maps were not enough to provide fine mapping of QTL and identification of candidate genes due to low polymorphism rate of traditional markers. Consequently, SNP markers with high abundance and even distribution in genome are more suitable for constructing ultra-dense genetic map. In this study, we developed a high-density genetic map of intraspecific Asiatic cotton, containing 3286 bin markers with the percentage of < 5 cM gaps more than 99%. Meanwhile, the collinearity between physical and genetic map was good except the five markers on Chr02 were matched to LG08 (Additional file 6: Table S6). Taken together, the quality this genetic map was good and subsequent QTL mapping of fiber traits was reliable.

In fact, an ultra-dense genetic linkage map is tremendously beneficial for assembling large and complex polyploidy genomes [32]. In the present study, we assembled 60 Block markers of 13 chrts into genetic map based on Illumina sequencing reads (Additional file 6: Table S6). These previously unanchored chrts were mapped on chromosome 02, 04, 05, 06, 07, 08, 11 and 12 respectively, which covered a total of 5.62 Mb. Eventually, the Asiatic cotton genome sequences will be updated, which will benefit the integrity of genome structure and detection of candidate genes for key agronomic traits.

Segregation distortion

Segregation distortion is common in most QTL mapping populations, and is common in cotton [33, 34]. Genetic causes of this phenomenon can be hybrid incompatibility and no-functional gamete formation, which related to species, population types and crosses. And sometimes marker types also lead to distortion [35-37]. Segregation distortion markers may link to genes or traits of interest [38-40]. Hence, filtering out these distortion markers may result in loss of important gene information. The present study included SDMs (p<0.05) for linkage map construction. Besides, we found the proportion of SDMs was relatively low (13.57%) and these markers were mainly derived from SXY No.1 alleles (72.42%), which showed that SXY No.1 plays a significant role in segregation distortion. Among 48 stable QTL identified in this study, only confidence interval of qUQL01.4 and qIFC13.1 located on segregation distortion regions (Figure 1), suggesting that segregation distortion markers were not linked to the majority of stable QTL.

QTL identification
In the present study, there were significant differences for fiber traits between parents and their progeny population showed transgressive segregation with huge differences, which indicated that both parents had diverse alleles for these fiber traits. Obviously, each of the parents contributed equally for fiber quality traits (72: 64). However, SXY No.1 contributed more favorable alleles than CSLZ for lint yield traits (30: 11), especially seed index (22: 5), which indicated parents with similar traits may confer different number of favorable QTL alleles to their offspring.

Among a total of 177 QTL for fiber traits detected in this study, 131 (74.0%) QTL were found in 47 clusters across 13 chromosomes. Many QTL belonged to same QTL cluster shared an overlapping region, which suggested that these overlapping regions may contain pleiotropic genes or multiple genes related to different fiber traits [41-43]. In this study, for example, Chr13-cluster-3 contained qFF13.2, qFNC13.1 and qLP13.1, sharing a nearly same confidence interval (Table 3), which indicated candidate genes in this region might control all three traits. Thus, candidate genes of clustered QTL may participate in the complex genetic network of fiber development via regulating multiple fiber traits [20, 44].

**Stable and common QTL**

Variance analysis in the present study showed that the environment had significant effect on eight key fiber traits. Consequently, QTL which were detected more than one environment were regarded as stable QTL. Our results indicated that 48 of 177 detected QTL occurred in at least two environments, including 38 QTL for fiber quality and 10 QTL for lint yield, most of which (72.9%) were within QTL clusters. These stable QTL and clusters may deserve high priorities for fine mapping and identification of candidate genes in the future.

In addition, we compared the stable QTL with QTL listed in the COTTONGEN [45], We employed sequences of flanking markers to do blast alignment in upland cotton genome [26], then the physical position of result sequences were found on COTTONGEN. In total, Twelve stable QTL (qUQL01.4, qUQL08.4, qUQL10.1, qUQL11.1, qUQL12.1, qUQL13.1, qIFC13.1, qLP12.1, qLP13.1, qSI07.2, qSI07.3 and qSI09.4) had been previously reported, while the rest were newly found in this study. These common QTL could be desirable alleles to improve fiber quality and lint yield by transferring into cotton cultivars.

**Conclusions**

This research was the first to use RIL population of *G. arboretum* and whole genome resequencing to map and identify the key fiber traits. The genetic map covered 1980.17 cM with an average distance of 0.61 cM between adjacent bin markers. A total of 48 stable QTL for 6 key fiber traits (UQL, FF, IFC, FNC, LP, SI) were found across 3 environments.

**Methods**

**Population construction**

*G. arboreum* cultivars SXY No.1 and CSLZ were crossed to develop a RIL mapping population. SXY No.1, the male parent was from Yellow River Region. CSLZ, the female parent came from Yangtze River Region. There were significant differences in multiple fiber-related traits between the two parents. The parents were crossed at Southwest University, Chongqing, China, in the summer of 2014. Single-seed descent was used from $F_{2:3}$ to $F_{2:8}$ to produce a RIL population with 189 lines. The RIL population was formed in the summer of 2018.

**Phenotypic data analysis**

Phenotyping of fiber-related traits including yield and quality was performed in three different environments (2018, 2019 in Chongqing, 2018 in Hainan). Yield component traits including lint percentage (LP, %) and seed index (SI, g/100 seeds) were tested. LP was calculated from lint weight divided by seed cotton weight. SI was measured by weighing 100 cotton seeds. As for fiber quality traits, fiber samples from whole population were evaluated with AFIS (Advanced Fiber Information System) at the Supervision Inspection and Testing Cotton Quality Center, Anyang, China. Data were collected on upper quartile length (UQL, mm), fiber fineness (FF, mtex), maturity ratio (MR), immature fiber content (IFC, %), fiber nep count (FNC, n·g$^{-1}$), fiber neps mean
size (FNMS, μm). SPSS version 21 (SPSS, Chicago, IL USA) was employed to analyze correlation coefficients among traits and analysis of variance (ANOVA). The distributions and probability density of these traits were illustrated intuitively in violin plot drawing by R x64 3.6.2.

**DNA isolation, library construction and high throughput sequencing**

Fresh young leaves of two parents and 189 RILs were collected for total genomic DNA extraction by a modified CTAB method [46]. The genomic DNA was randomly cleaved into 200-500 bp fragments by ultrasonic method, then DNA library was constructed by terminal repair, addition of polyA-tails and paired-end adaptors, purification and PCR amplification of these small DNA fragments. Eventually paired-end sequencing was performed with an Illumina HiSeqTM-2500 platform (Illumina, Inc., San Diego, CA, USA) at the Biomarker Technology Co. Ltd. (Beijing, China). In order to ensure the accuracy of data, raw reads were filtered to get clean reads for subsequent data analysis. The main steps of data filtering are as follows: removing the adaptors; removing reads with >10% unidentified nucleotides (N); and removing the low-quality reads with >50% bases of Q \text{phred} ≤ 10.

**SNP detection**

The clean reads were aligned to the Asian cotton reference genome [16] using BWA-MEM algorithm (version 0.7.15) [47] for genomic variation analysis. The SAMtools software (version 1.6) [48] was used to convert the sequence alignment files (SAM) containing mapping results into binary BAM files, and to sort, merge and count alignment efficiency of BAM files. Subsequently, the SNP detection of parents and RILs was implemented by GATK (Genome Analysis Toolkit) software (version 3.8) [49] including Indel realignment, base recalibration, variant calling and final SNP filtration. The SNP variation results were displayed in VCF file format and input to SnpEff (SNP effect) software (version 1.9.6) [50] to annotate SNP variation. The parameters of the above software were all default. The polymorphic SNPs which were identified between the parents were used to genotype calling. Before genotype calling, unqualified SNPs were filtered out according to following criteria: the markers of intergenic region; the markers that were homozygous and inconsistent in two patents; the read depth of parental markers was less than 4-fold.

**Genotyping and bin map construction**

As for genotype calling, the missing genotypes were imputed using k-nearest neighbor algorithm [51]. A window size of 15 SNPs, with a step size of 1 SNP, was a sliding unit for chromosome scan. When the aa: bb allele ratio was 11:4 or higher in the sliding window, the segregation pattern was aa. When the aa: bb allele ratio was 4:11 or lower in the sliding window, the segregation pattern was bb. Other cases with intermediate ratios, ab pattern was used for genotype filling and correction. The SNPs whose genotypes were aa ’ bb were employed for genetic map construction.

Recombination breakpoint was assumed when the sliding window moves on the chromosome until genotype changes. The SNP alleles between the recombination breakpoints was designated as a recombination bin. Unqualified bins were excluded if the length of bins was less than 20 kb, and bins showing significant deviation from the 1:1 were also filtered out through chi-square test (p<0.01). Bin markers were used to construct linkage groups with Highmap software [52]. The SMOOTH [53] was applied to correct errors according to genotypes of parental contribution. The Kosambi mapping function was used to estimate genetic map distances [54].

**QTL analysis**

QTL identification was performed with Map QTL 6.0 [55]. A threshold of log of odds ratio (LOD) was set as 2.0 [56]. Positive additive effects indicated favorable alleles derived from SXY No.1 whereas negative additive effects indicated favorable alleles derived from CSLZ. QTL identified in two or three environments were considered to be potential stable QTL in this study. MapChart 2.2 [57] was used for presentation of genetic map and QTL. The nomenclature of QTL was: q + trait abbreviation + chromosome number + QTL number. QTL for the same trait across different environments were considered in the same QTL region when their confidence intervals overlapped.

Segregation distorted regions (SDRs): the red labeled part of chromosomes
Table 2 Stable QTL for fiber quality and yield traits detected in the SXY No.1 ´ CSLZ RIL population
| Traita | QTL | Chr | Envb | Flanking markers | Location | Nearest locus | LOD | Additivec (%) | PVEd (%) |
|--------|-----|-----|------|------------------|----------|--------------|-----|---------------|----------|
| UQL   | qUQL01.4 | Chr01 | 2018CQ | Block851 Block1091 | 108.819 | Block921 | 4.38 | 0.59 | 10.6 |
| | | Chr01 | 2018HN | Block1052 Block1091 | 127.878 | Block1054 | 2.16 | 0.34 | 5.4 |
| | qUQL03.1 | Chr03 | 2018HN | Block1996 Block2031 | 18.02 | Block2001 | 2.2 | 0.34 | 5.5 |
| | | Chr03 | 2019CQ | Block1972 Block2039 | 20.522 | Block2031 | 4.43 | 0.49 | 10.8 |
| | qUQL03.3 | Chr03 | 2018CQ | Block2759 Block2836 | 79.659 | Block2784 | 3.49 | 0.52 | 8.6 |
| | | Chr03 | 2018HN | Block2775 Block2853 | 81.813 | Block2811 | 3.26 | 0.41 | 8 |
| | qUQL04.3 | Chr04 | 2018HN | Block3871 Block3866 | 63.163 | Block3868 | 2.35 | 0.43 | 5.9 |
| | | Chr04 | 2018HN | Block3839 Block3792 | 69.886 | Block3817 | 2.74 | 0.38 | 6.8 |
| | | Chr04 | 2019CQ | Block3831 Block3789 | 70.152 | Block3816 | 3.06 | 0.42 | 7.6 |
| | qUQL06.2 | Chr04 | 2018CQ | Block6948 Block7004 | 82.399 | Block6957 | 3.3 | 0.51 | 8.1 |
| | | Chr04 | 2018HN | Block6948 Block7004 | 83.037 | Block6971 | 3.28 | 0.42 | 8.1 |
| | qUQL06.3 | Chr04 | 2018CQ | Block7105 Block7130 | 102.688 | Block7119 | 4.26 | 0.47 | 10.4 |
| | | Chr04 | 2018HN | Block7105 Block7130 | 102.688 | Block7119 | 4.26 | 0.47 | 10.4 |
| | qUQL07.1 | Chr07 | 2018HN | Block8438 Block8484 | 119.227 | Block8465 | 3.2 | -0.41 | 7.9 |
| | | Chr07 | 2019CQ | Block8438 Block8484 | 120.708 | Block8471 | 2.26 | -0.36 | 5.7 |
| | qUQL08.3 | Chr08 | 2018CQ | Block9236 Block9283 | 47.507 | Block9275 | 3.62 | 0.53 | 8.9 |
| | | Chr08 | 2018HN | Block9276 Block9290 | 49.679 | Block9283 | 2.69 | 0.37 | 6.7 |
| | qUQL08.4 | Chr08 | 2018CQ | Block9303 Block9348 | 56.145 | Block9325 | 3.49 | 0.52 | 8.6 |
| | | Chr08 | 2018HN | Block9303 Block9348 | 56.145 | Block9323 | 2.33 | 0.35 | 5.8 |
| | qUQL10.1 | Chr10 | 2018HN | Block11119 Block1185 | 43.535 | Block11177 | 2.88 | -0.39 | 7.1 |
| | | Chr10 | 2019CQ | Block11150 Block1185 | 43.535 | Block11177 | 2.66 | -0.39 | 6.6 |
| | qUQL11.1 | Chr11 | 2018HN | Block12879 Block12918 | 68.27 | Block12884 | 6.32 | -0.56 | 15 |
| | | Chr11 | 2019CQ | Block12864 Block12917 | 65.041 | Block12867 | 4.63 | -0.50 | 11.2 |
| | qUQL11.3 | Chr11 | 2018CQ | Block13419 Block13464 | 145.656 | Block13440 | 3.11 | 0.50 | 7.7 |
| | | Chr11 | 2018HN | Block13415 Block13453 | 145.656 | Block13440 | 4.35 | 0.47 | 10.6 |
| | | Chr11 | 2019CQ | Block13419 Block13453 | 145.656 | Block13440 | 3.12 | 0.42 | 7.7 |
| | qUQL12.1 | Chr12 | 2018CQ | Block13627 Block13657 | 0.691 | Block13650 | 2.73 | -0.46 | 6.8 |
| | | Chr12 | 2018HN | Block13627 Block13657 | 0.691 | Block13650 | 2.97 | -0.39 | 7.4 |
| | qUQL12.2 | Chr12 | 2018CQ | Block13914 Block13935 | 53.904 | Block13924 | 2.88 | 0.48 | 7.1 |
| | | Chr12 | 2019CQ | Block13896 Block13929 | 53.264 | Block13918 | 2.72 | 0.39 | 6.8 |
| | qUQL13.1 | Chr13 | 2018HN | Block14798 Block14726 | 77.034 | Block14691 | 2.4 | -0.36 | 6 |
| | | Chr13 | 2019CQ | Block14798 Block14726 | 79.052 | Block14716 | 3.13 | -0.42 | 7.7 |
| FF   | qFF02.1 | Chr02 | 2018CQ | Block15854 Block1683 | 0 | Block15854 | 3.92 | -7.02 | 9.6 |
| | | Chr02 | 2018HN | Block15854 Block1673 | 4.383 | Block15839 | 2.67 | -5.32 | 6.6 |
| Trait<sup>a</sup> | QTL | Chr | Environ<sup>b</sup> | Flanking markers | Location | Nearest locus | LOD | Additive<sup>c</sup> | PVE<sup>d</sup> (%) |
|------------|-----|-----|-----------------|-----------------|----------|--------------|-----|----------------|------------------|
| qFF02.3    | Chr02 | 2019CQ | Block15854 | Block1326 | 0 | Block15854 | 2.4 | -6.37 | 6 |
| qFF02.4    | Chr02 | 2018CQ | Block1371 | Block1341 | 69.181 | Block1353 | 2.99 | -6.34 | 7.4 |
| qFF02.5    | Chr02 | 2018CQ | Block1327 | Block1312 | 83.054 | Block1325 | 2.66 | -5.90 | 6.6 |
| qFF03.1    | Chr03 | 2018CQ | Block1873 | Block1911 | 12.826 | Block1891 | 2.35 | -5.51 | 5.9 |
| qFF03.4    | Chr03 | 2018HN | Block2742 | Block2784 | 78.329 | Block2761 | 4.22 | -6.65 | 10.3 |
| qFF05.1    | Chr05 | 2018CQ | Block5046 | Block5142 | 101.827 | Block5100 | 2.6 | 5.80 | 6.5 |
| qFF05.2    | Chr05 | 2018CQ | Block5260 | Block5305 | 125.289 | Block5293 | 3.02 | 6.28 | 7.5 |
| qFF05.3    | Chr05 | 2018CQ | Block5578 | Block5607 | 154.151 | Block5602 | 2.29 | 5.54 | 5.7 |
| qFF05.4    | Chr05 | 2018CQ | Block6387 | Block6563 | 202.929 | Block6481 | 4.8 | 7.76 | 11.6 |
| qFF06.1    | Chr06 | 2018CQ | Block6752 | Block6808 | 34.849 | Block6794 | 7.59 | -9.85 | 17.7 |
| qFF07.1    | Chr07 | 2018CQ | Block7822 | Block7897 | 45.016 | Block7861 | 4.56 | 7.62 | 11.1 |
| qFF07.2    | Chr07 | 2018CQ | Block8141 | Block8197 | 84.061 | Block8146 | 3.7 | 6.84 | 9.1 |
| qFF10.2    | Chr10 | 2018CQ | Block11105 | Block1132 | 35.956 | Block11126 | 4.49 | 7.61 | 10.9 |
| qFF11.1    | Chr11 | 2018CQ | Block12040 | Block1238 | 7.467 | Block12164 | 2.89 | 6.09 | 7.2 |

<sup>a</sup> Trait: Characteristic trait studied.
<sup>b</sup> Environ: Environment under which the trait was measured.
<sup>c</sup> Additive: Additive effect of the QTL.
<sup>d</sup> PVE: Percentage of variance explained.
| Trait<sup>a</sup> | QTL | Chr  | Env<sup>b</sup> | Flanking markers | Location | Nearest locus | LOD | Additive<sup>c</sup> | PVE<sup>d</sup> (%) |
|----------------|-----|------|----------------|-----------------|----------|--------------|-----|----------------|------------------|
|                |     |      |                |                 |          |              |     |                 |                  |
|                |     |      |                |                 |          |              |     |                 |                  |
| IFC            | qFC01.1 | Chr01 | 2018HN | Block567|Block588 | 47.508 | Block572 | 2.83 | 10.02 | 7                  |
|                | qFC01.1 | Chr01 | 2018HN | Block567|Block588 | 47.508 | Block572 | 2.3  | 3.74  | 5.7                 |
| FNC            | qFC07.1 | Chr07 | 2018CQ | Block7702|Block7768 | 33.788 | Block7761 | 5.87 | -4.44 | 14                 |
|                | qFC07.1 | Chr07 | 2018HN | Block7702|Block7822 | 27.05  | Block7713 | 3.11 | -10.50 | 7.7                |
|                | qFC07.1 | Chr07 | 2018CQ | Block7734|Block7774 | 32.671 | Block7751 | 3.33 | -4.48 | 8.2                |
|                | qFC07.5 | Chr07 | 2018CQ | Block8570|Block8622 | 140.245| Block8577 | 2.94 | -3.21 | 7.3                |
|                | qFC11.1 | Chr11 | 2018CQ | Block12927|Block13045 | 78.82  | Block13043 | 2.89 | -10.09 | 7.2                |
|                | qFC11.1 | Chr11 | 2019CQ | Block12927|Block12949 | 73.738 | Block12941 | 3.28 | -4.42 | 8.1                |
| LP             | qLP11.1 | Chr11 | 2018CQ | Block13185|Block13198 | 113.372| Block13189 | 2.08 | -0.70 | 5.2                |
|                | qLP11.1 | Chr11 | 2019CQ | Block13171|Block13198 | 112.571| Block13177 | 2.19 | -0.72 | 5.5                |
|                | qLP12.1 | Chr12 | 2018CQ | Block13884|Block13890 | 47.642 | Block13888 | 2.09 | -0.71 | 5.2                |
|                | qLP12.1 | Chr12 | 2019CQ | Block13896|Block13913 | 50.851 | Block13901 | 2.44 | -0.76 | 6.1                |
|                | qLP13.1 | Chr13 | 2018CQ | Block14716|Block14749 | 84.269 | Block14731 | 6.47 | 1.23  | 15.3               |
|                | qLP13.1 | Chr13 | 2018HN | Block14716|Block14749 | 86.028 | Block14742 | 10.68 | 1.38  | 24                 |
|                | qLP13.1 | Chr13 | 2019CQ | Block14726|Block14740 | 84.269 | Block14731 | 18.2 | 1.90  | 37.4               |
| SI             | qSI01.3 | Chr01 | 2018CQ | Block851 |Block1004 | 109.354| Block923  | 3.8  | 0.18  | 9.3                |
|                | qSI01.3 | Chr01 | 2018HN | Block851 |Block921  | 105.324| Block901  | 6.15 | 0.25  | 14.6               |
|                | qSI04.1 | Chr04 | 2018CQ | Block3806|Block3787 | 74.972 | Block3790 | 2.29 | 0.14  | 5.7                |
|                | qSI04.1 | Chr04 | 2018HN | Block3801|Block3789 | 71.585 | Block3797 | 2.04 | 0.15  | 5.1                |
|                | qSI04.1 | Chr04 | 2019CQ | Block3787|Block3762 | 81.335 | Block3767 | 2.24 | 0.17  | 5.6                |
|                | qSI07.1 | Chr07 | 2018CQ | Block7758|Block7899 | 47.68  | Block7882 | 3.04 | 0.16  | 7.5                |
|                | qSI07.1 | Chr07 | 2018HN | Block7764|Block7835 | 36.96  | Block7822 | 3.17 | 0.18  | 7.8                |
|                | qSI07.2 | Chr07 | 2018CQ | Block8153|Block8423 | 92.432 | Block8193 | 5.58 | 0.21  | 13.4               |
### Table 3: QTL Clusters for Fiber Quality and Yield Traits Identified Across Three Environments in the SXY No.1 ´ CSLZ RIL Population

| Trait | QTL | Chr   | Env<sup>b</sup> | Flanking markers | Location | Nearest locus | LOD  | Additive<sup>c</sup> | PVE<sup>d</sup> (%) |
|-------|-----|-------|-----------------|------------------|----------|---------------|------|---------------------|--------------------|
|       | Chr07 | 2018HN | Block8167Block8191 | 90.262 | Block8188 | 4.17 | 0.21 | 10.2 |
| qSI07.3 | Chr07 | 2018CQ | Block8570Block8600 | 140.721 | Block8582 | 7.59 | 0.24 | 17.7 |
|       | Chr07 | 2018HN | Block8530Block8672 | 136.396 | Block8552 | 3.27 | 0.19 | 8.1  |
| qSI08.3 | Chr08 | 2018CQ | Block9712Block9737 | 138.202 | Block9726 | 6.15 | 0.22 | 14.6 |
|       | Chr08 | 2018HN | Block9712Block9737 | 138.202 | Block9726 | 6.9  | 0.26 | 16.3 |
| qSI09.4 | Chr09 | 2018HN | Block10404Block10571 | 114.878 | Block10502 | 3.68 | 0.20 | 9    |
|       | Chr09 | 2019CQ | Block10404Block10597 | 114.878 | Block10502 | 4.77 | 0.25 | 11.5 |

<sup>a</sup>UQL, upper quartile length; FF, fiber fineness; MR, maturity ratio; IFC, immature fiber content; FNC, fiber neps count; FNMS, fiber neps mean size; LP, lint percentage; SI, seed index.

<sup>b</sup>2018CQ, 2018 at Chongqing; 2018HN, 2018 at Hainan; 2019CQ, 2019 at Chongqing.

<sup>c</sup>Positive additive effects indicated that SXY No.1 alleles increased the phenotypic value, negative additive effects suggested that CSLZ alleles increased the phenotypic value.

<sup>d</sup>Phenotypic variance explained.
| Cluster/hotpot | Flanking markers | QTL |
|---------------|-----------------|-----|
| Chr01-cluster-1 | Block544Block694 | qFNC01.1\#, qMR01.1, qSI01.1 |
| Chr01-cluster-2 | Block692Block804 | qUQL01.3, qMR01.2 |
| Chr02-cluster-1 | Block15854Block1665 | qUQL02.1, qFF02.1\#, qUQL02.2 |
| Chr02-cluster-2 | Block1215Block1105 | qFNC02.1, qFF02.5\#, qMR02.1, qIFC02.1 |
| Chr03-cluster-1 | Block1873Block2084 | qFF03.1\#, qLP03.1, qUQL03.1\# |
| Chr03-cluster-2 | Block2206Block2448 | qFF03.2, qLP03.2, qUQL03.2 |
| Chr03-cluster-3 | Block2742Block2904 | qFF03.4\#, qUQL03.3\#, qIFC03.1, qMR03.1 |
| Chr04-cluster-1 | Block4484Block4159 | qFF04.1, qUQL04.1 |
| Chr04-cluster-2 | Block4179Block3927 | qFNC04.1, qIFC04.1, qFNMS04.1 |
| Chr04-cluster-3 | Block3950Block3905 | qIFC04.2, qUQL04.2 |
| Chr04-cluster-4 | Block3871Block3762 | qUQL04.3\#, qSI04.1\# |
| Chr05-cluster-1 | Block5046Block5196 | qUQL05.1, qFF05.2\# |
| Chr05-cluster-2 | Block5206Block5305 | qUQL05.2, qFF05.2\# |
| Chr05-cluster-3 | Block5509Block5695 | qSI05.3, qFF05.3\# |
| Chr06-cluster-1 | Block6728Block6808 | qMR06.1, qFNC06.1, qFF06.1\# |
| Chr06-cluster-2 | Block6948Block7038 | qUQL06.2\#, qSI06.1 |
| Chr06-cluster-3 | Block7105Block7178 | qUQL06.3\#, qSI06.2 |
| Chr07-cluster-1 | Block7603Block7693 | qLP07.1, qFNMS07.1 |
| Chr07-cluster-2 | Block7758Block7899 | qSI07.1\#, qFF07.1\#, qMR07.1, qFNC07.2 |
| Chr07-cluster-3 | Block8141Block8197 | qFF07.2\#, qMR07.2, qLP07.2 |
| Chr07-cluster-4 | Block8153Block8423 | qFNC07.4, qSI07.2\# |
| Chr07-cluster-5 | Block8570Block8622 | qFNC07.5\#, qMR07.3 |
| Chr07-cluster-6 | Block8679Block8780 | qFF07.3, qIFC07.1 |
| Chr07-cluster-7 | Block8780Block8910 | qFNMS07.2, qSI07.4, qIFC07.2 |
| Chr07-cluster-8 | Block8956Block9034 | qIFC07.3, qFF07.4 |
| Chr08-cluster-1 | Block9236Block9364 | qUQL08.3\#, qFF08.1, qLP08.1, qUQL08.4\#, qFNMS08.1 |
| Chr08-cluster-2 | Block9373Block9434 | qUQL08.5, qSI08.1 |
| Chr08-cluster-3 | Block9459Block9520 | qFNMS08.2, qUQL08.6 |
| Chr08-cluster-4 | Block9645Block9737 | qFF08.2, qMR08.1, qSI08.3\# |
| Clusters/hotspots | Flanking markers | QTL \(^b\) |
|------------------|-----------------|--------------|
| Chr09-cluster-1  | Block9905Block9963 | qFF09.1, qMR09.1, qIFC09.1, qSI09.1 |
| Chr09-cluster-2  | Block10064Block10120 | qSI09.2, qMR09.2, qIFC09.2 |
| Chr09-cluster-3  | Block10222Block10353 | qFF09.2, qSI09.3 |
| Chr09-cluster-4  | Block10624Block10717 | qFNMS09.1, qLP09.2 |
| Chr10-cluster-1  | Block10924Block10999 | qFF10.1, qMR10.1, qIFC10.1 |
| Chr10-cluster-2  | Block11105Block11196 | qFF10.2\*, qUQL10.1\* |
| Chr10-cluster-3  | Block11223Block11374 | qSI10.1, qLP10.1 |
| Chr10-cluster-4  | Block11676Block11763 | qFF10.3, qMR10.2, qIFC10.2, qFNMS10.1, qFNC10.1 |
| Chr10-cluster-5  | Block11817Block11912 | qLP10.2, qUQL10.2 |
| Chr11-cluster-1  | Block12545Block12779 | qFF11.2\#, qSI11.1 |
| Chr11-cluster-2  | Block12864Block13045 | qUQL11.1\#, qFNC11.1\#, qIFC11.1 |
| Chr11-cluster-3  | Block13171Block13226 | qLP11.1\#, qFNMS11.1 |
| Chr12-cluster-1  | Block13839Block13935 | qFF12.1, qMR12.1, qIFC12.1, qLP12.1\#, qUQL12.2\#, qSI12.1 |
| Chr13-cluster-1  | Block15471Block15155 | qFF13.1, qMR13.1, qIFC13.1\#, qFNMS13.1 |
| Chr13-cluster-2  | Block14801Block14726 | qSI13.2, qUQL13.1\* |
| Chr13-cluster-3  | Block14716Block14676 | qFF13.2\#, qFNC13.1, qLP13.1\# |

\*indicates stable QTL identified across two environments. \#indicate stable QTL identified across three environments.

\(^b\)UQL, upper quartile length; FF, fiber fineness; MR, maturity ratio; IFC, immature fiber content; FNC, fiber neps count; FNMS, fiber neps mean size; LP, lint percentage; SI, seed index.

**Abbreviations**

Chr: Chromosomes; FF: fiber fineness; FNC: fiber neps count; FNMS: fiber neps mean size; GBS: Genotyping by sequencing; IFC: immature fiber content; LP: lint percentage; LOD: Log of Odds; MAS: marker assisted selection; MR: maturity ratio; NGS: next generation sequencing; PVE: Phenotypic variance explained; QTL: Quantitative trait loci; RAD-seq: Restriction-site associated DNA sequencing; RIL, Recombinant inbred line; SI: seed index; SNP: Single nucleotide polymorphism; SDM, Segregation distortion marker; SDR: Segregation distortion region; UQL: upper quartile length;

**Declarations**

**Ethics approval and consent to participate**

Not applicable.

**Consent for publication**

Not applicable.

**Availability of data and materials**
Sequencing data related to this study has been uploaded to NCBI SRA database, which can be accessed through series of SRA numbers PRJNA649393.

Competing interests
The authors declare that they have no competing interests.

Funding
This study was partially funded by the Genetically Modified Organisms Breeding Major Project of China (2016ZX08005005-001 to Y.H.X.), the National Natural Science Foundation of China (31571582 and 31871539 to Y.H.X.) and Chongqing Science and Technology Bureau (cstc2018scx-msybX0382 to Y.H.X.).

Authors’ contributions
YHX conceived the study, participated in its design and modified the manuscript. YHL contributed to data analysis and manuscript writing. TM, LFR, JYZ collected the data from the field; CNW, ZMZ, AML contributed to population construction. All authors have read and approved the final manuscript.

Acknowledgments
Not applicable.

References
1. Wendel JF, Brubaker C, Alvarez I, Cronn R, Stewart JM: Evolution and Natural History of The Cotton Genus. In: Paterson, AH (ed) Genetics and genomics of cotton, 2009, plant genetics and genomics: crops and models 3:pp 3-22.
2. Beasley JO: The origin of American tetraploid Gossypium species. Am Nat 1940, 74:285-286.
3. Chen ZJ, Scheffler BE, Dennis E, Triplett BA, Zhang T, Guo W, Chen X, Stelly DM, Rabinowicz PD, Town CD et al: Toward sequencing cotton (Gossypium) genomes. Plant Physiol 2007, 145(4):1303-1310.
4. Lubbers EL, Chee PW: The worldwide gene pool of G. hirsutum and its improvement. In: Paterson, AH (ed) Genetics and genomics of cotton, 2009, plant genetics and genomics: crops and models 3:pp 23-52.
5. Kulkarni VN, Khadi, B.M., Maralappanavar, M.S., Deshapande, L.A., and Narayanan S.S.: The Worldwide Gene Pools of Gossypium arboreum L. and G. herbaceum L., and Their Improvement. In: Paterson, AH (ed) Genetics and genomics of cotton, 2009, plant genetics and genomics: crops and models 3:pp 69-97.
6. Liu F, Zhou ZL, Wang CY, Wang YH, Cai XY, Wang XX, Zhang ZS, Wang KB: Genetic diversity and relationship analysis of Gossypium arboreum accessions. Genet Mol Res 2015, 14(4):14522-14529.
7. Maqbool A, Abbas W, Rao AQ, Irfan M, Zahur M, Bakhsh A, Riazuddin S, Husnain T: Gossypium arboreum GHSP26 enhances drought tolerance in Gossypium hirsutum. Biotechnol Prog 2010, 26(1):21-25.
8. Stanton MA, Stewart JMD, Tugwell NPJGR, Evolution C: Evaluation of Gossypium arboreum L. Germplasm for resistance to thrips. 1992, 39(2):89-95.
9. Mehetre SS, Ahir AR, Gawande VL, Patil VR, Mokate AS: Induced polyploidy in Gossypium: A tool to overcome interspecific incompatibility of cultivated tetraploid and diploid cottons. Curr Sci India 2003, 84(12):1510-1512.
10. Tahir MS, Khan NUI, Sajid-Ur-Rehman: Development of an Interspecific Hybrid (Triploid) by Crossing Gossypium hirsutum and G. arboreum. Cytologia 2011, 76(2):193-199.
11. Kebede H, Burow G, Dani RG, Allen RD: A-genome cotton as a source of genetic variability for Upland cotton (Gossypium hirsutum). Genet Resour Crop Ev 2007, 54(4):885-895.
12. Kantartzi SK, Ulloa M, Sacks E, Stewart JM: Assessing genetic diversity in Gossypium arboreum L. cultivars using genomic and EST-derived microsatellites. Genetica 2009, 136(1):141-147.
13. Chen Z, Wang B, Dong X, Liu H, Ren L, Chen J, Hauck A, Song W, Lai J: An ultra-high density bin-map for rapid QTL mapping for tassel and ear architecture in a large F(2) maize population. BMC Genomics 2014, 15:433.

14. Zhu MS, Liu DL, Liu WG, Li D, Liao YL, Li JH, Fu CY, Fu FH, Huang HJ, Zeng XQ et al: QTL mapping using an ultra-high-density SNP map reveals a major locus for grain yield in an elite rice restorer R998. Sci Rep-Uk 2017, 7.

15. Thyssen GN, Jenkins JN, McCarty JC, Zeng L, Campbell BT, Delhom CD, Islam MS, Li P, Jones DC, Condon BD et al: Whole genome sequencing of a MAGIC population identifies genomic loci and candidate genes for major fiber quality traits in upland cotton (Gossypium hirsutum L.). Theor Appl Genet 2019, 132(4):989-999.

16. Du X, Huang G, He S, Yang Z, Sun G, Ma X, Li N, Zhang X, Sun J, Liu M et al: Resequencing of 243 diploid accessions based on an updated A genome identifies the genetic basis of key agronomic traits. Nat Genet 2018, 50(6):796-802.

17. Ma X, Wang Z, Li W, Zhang Y, Zhou X, Liu Y, Ren Z, Pei X, Zhou K, Zhang W et al: Resequencing core accessions of a pedigree identifies derivation of genomic segments and key agronomic trait loci during cotton improvement. Plant Biotechnol J 2019, 17(4):762-775.

18. Jia X, Pang C, Wei H, Wang H, Ma Q, Yang J, Cheng S, Su J, Fan S, Song M et al: High-density linkage map construction and QTL analysis for earliness-related traits in Gossypium hirsutum L. BMC Genomics 2016, 17(1):909.

19. Wang HT, Jin X, Zhang BB, Chen C, Lin ZX: Enrichment of an intraspecific genetic map of upland cotton by developing markers using parental RAD sequencing. DNA Res 2015, 22(2):147-160.

20. Ali I, Teng Z, Bai Y, Yang Q, Hao Y, Hou J, Jia Y, Tian L, Liu X, Tan Z et al: A high density SLAF-SNP genetic map and QTL detection for fibre quality traits in Gossypium hirsutum. BMC Genomics 2018, 19(1):879.

21. Wang W, Sun Y, Yang P, Cai X, Yang L, Ma J, Ou Y, Liu T, Ali I, Liu D et al: A high density SLAF-seq SNP genetic map and QTL for seed size, oil and protein content in upland cotton. BMC Genomics 2019, 20(1):599.

22. Wang F, Zhang J, Chen Y, Zhang C, Gong J, Song Z, Zhou J, Wang J, Zhao C, Jiao M et al: Identification of candidate genes for key fibre-related QTLs and derivation of favourable alleles in Gossypium hirsutum recombinant inbred lines with G. barbadense introgressions. Plant Biotechnol J 2020, 18(3):707-720.

23. Diouf L, Magwanga RO, Gong WF, He SP, Pan ZE, Jia YH, Kirungu JN, Du XM: QTL Mapping of Fiber Quality and Yield-Related Traits in an Intra-Specific Upland Cotton Using Genotype by Sequencing (GBS). International Journal of Molecular Sciences 2018, 19(2).

24. Fan LP, Wang LP, Wang XY, Zhang HY, Zhu YF, Guo JY, Gao WW, Geng HW, Chen QJ, Qu YY: 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 2019, 20(1):599.

25. Paterson AH, Wendel JF, Gundlach H, Guo H, Jenkins J, Jin D, Llewellyn D, Showmaker KC, Shu S, Udall J et al: Repeated polyploidization of Gossypium genomes and the evolution of spinnable cotton fibres. Nature 2012, 492(7429):423-427.

26. Zhang T, Hu Y, Jiang W, Fang L, Guan X, Chen J, Zhang J, Saski CA, Scheffler BE, Stelly DM 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.

27. Wang M, Tu L, Yuan D, Zhu, Shen C, Li J, Liu F, Pei L, Wang P, Zhao G et al: Reference genome sequences of two cultivated allotetraploid cottons, Gossypium hirsutum and Gossypium barbadense. Nat Genet 2019, 51(2):224-229.

28. Ma XX, Zhou BL, Lu YH, Guo WZ, Zhang TZ: Simple sequence repeat genetic linkage maps of A-genome diploid cotton (Gossypium arboreum). J Integr Plant Biol 2008, 50(4):491-502.

29. Badigannavar A, Myers GO: Construction of Genetic Linkage Map and QTL Analysis for Fiber Traits in Diploid Cotton (Gossypium arboreum x Gossypium herbaceum). J Cotton Sci 2015, 19(1):15-26.

30. Desai A, Chee PW, Rong J, May OX, Paterson AH: Chromosome structural changes in diploid and tetraploid A genomes of Gossypium. Genome 2006, 49(4):336-345.

31. Li C, Dong Y, Zhao T, Li L, Li C, Yu E, Mei L, Daud MK, He Q, Chen J et al: Genome-Wide SNP Linkage Mapping and QTL Analysis for Fiber Quality and Yield Traits in the Upland Cotton Recombinant Inbred Lines Population. Front Plant Sci 2016, 7:1356.
32. Wang S, Chen J, Zhang W, Hu Y, Chang L, Fang L, Wang Q, Lv F, Wu H, Si Z et al: Sequence-based ultra-dense genetic and physical maps reveal structural variations of allopolyploid cotton genomes. Genome Biol 2015, 16:108.
33. Liu X, Teng Z, Wang J, Wu T, Zhang Z, Deng X, Fang X, Tan Z, Ali I, Liu D et al: Enriching an intraspecific genetic map and identifying QTL for fiber quality and yield component traits across multiple environments in Upland cotton (Gossypium hirsutum L.). Mol Genet Genomics 2017, 292(6):1281-1306.
34. Zhang L, Wang S, Li H, Deng Q, Zheng A, Li S, Li P, Li Z, Wang J: Effects of missing marker and segregation distortion on QTL mapping in F2 populations. Theor Appl Genet 2010, 121(6):1071-1082.
35. Zhang ZS, Hu MC, Zhang J, Liu DJ, Zheng J, Zhang K, Wang W, Wan Q: Construction of a comprehensive PCR-based marker linkage map and QTL mapping for fiber quality traits in upland cotton (Gossypium hirsutum L.). Mol Breeding 2009, 24(1):49-61.
36. Xu Y, Zhu L, Xiao J, Huang N, McCouch SR: Chromosomal regions associated with segregation distortion of molecular markers in F2, backcross, doubled haploid, and recombinant inbred populations in rice (Oryza sativa L.). Mol Gen Genet 1997, 253(5):535-545.
37. Takumi S, Motomura Y, lehisa JCM, Kobayashi F: Segregation distortion caused by weak hybrid necrosis in recombinant inbred lines of common wheat. Genetics 2013, 141(10-12):463-470.
38. Zhou W, Tang Z, Hou J, Hu N, Yin T: Genetic Map Construction and Detection of Genetic Loci Underlying Segregation Distortion in an Intraspecific Cross of Populus deltoides. PLoS One 2015, 10(5):e0126077.
39. Yin TM, DiFazio SP, Gunter LE, Riemenschneider D, Tuskan GA: Large-scale heterospecific segregation distortion in Populus revealed by a dense genetic map. Theor Appl Genet 2004, 109(3):451-463.
40. Cervera MT, Storme V, Ivens B, Gusmao J, Liu BH, Hostyn V, Van Slycken J, Van Montagu M, Boerjan W: Dense genetic linkage maps of three Populus species (Populus deltoides, P. nigra and P. trichocarpa) based on AFLP and microsatellite markers. Genetics 2001, 158(2):787-809.
41. Li SQ, Liu AY, Kong LL, Gong JW, Li JW, Gong WK, Lu QW, Li PT, Ge Q, Shang HH et al: QTL mapping and genetic effect of chromosome segment substitution lines with excellent fiber quality from Gossypium hirsutum x Gossypium barbadense. Mol Genet Genomics 2019, 294(5):1123-1136.
42. Abdelraheem A, Liu F, Song M, Zhang JF: A meta-analysis of quantitative trait loci for abiotic and biotic stress resistance in tetraploid cotton. Mol Genet Genomics 2017, 292(6):1221-1235.
43. Said JI, Lin Z, Zhang X, Song M, Zhang J: A comprehensive meta QTL analysis for fiber quality, yield related and morphological traits, drought tolerance, and disease resistance in tetraploid cotton. BMC Genomics 2013, 14:776.
44. Lacape JM, Llewellyn D, Jacobs J, Arioli T, Becker D, Calhoun S, Al-Ghazi Y, Liu S, Palai O, Georges S et al: Meta-analysis of cotton fiber quality QTLs across diverse environments in a Gossypium hirsutum x G. barbadense RIL population. BMC Plant Biol 2010, 10:132.
45. Yu J, Jung S, Cheng CH, Ficklin SP, Lee T, Zheng P, Jones D, Percy RG, Main D: CottonGen: a genomics, genetics and breeding database for cotton research. Nucleic Acids Res 2014, 42(Database issue):D1229-1236.
46. Zhang ZS, Xiao YH, Luo M, Li XB, Luo XY, Hou L, Li DM, Pei Y: Construction of a genetic linkage map and QTL analysis of fiber-related traits in upland cotton (Gossypium hirsutum L.). Euphytica 2005, 144(1-2):91-99.
47. Li H, Durbin R: Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009, 25(14):1754-1760.
48. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S: The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009, 25(16):2078-2079.
49. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M et al: The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 2010, 20(9):1297-1303.
50. Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM: A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin) 2012, 6(2):80-92.
51. Huang X, Zhao Y, Wei X, Li C, Wang A, Zhao Q, Li W, Guo Y, Deng L, Zhu C et al: Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nat Genet* 2011, 44(1):32-39.

52. Liu D, Ma C, Hong W, Huang L, Liu M, Liu H, Zeng H, Deng D, Xin H, Song J et al: Construction and analysis of high-density linkage map using high-throughput sequencing data. *PLoS One* 2014, 9(6):e98855.

53. van Os H, Stam P, Visser RG, van Eck HJ: SMOOTH: a statistical method for successful removal of genotyping errors from high-density genetic linkage data. *Theor Appl Genet* 2005, 112(1):187-194.

54. Kosambi DD: The estimation of map distances from recombination values. *Ann Eugenic* 1943, 12:172-175.

55. J.W. VO: MapQTL® 6.0. Software for the mapping of quantitative trait loci in experimental populations of diploid species. In. Wageningen. The Netherlands: Kyazma BV; 2009.

56. Lander E, Kruglyak L: Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 1995, 11(3):241-247.

57. Voorrips RE: MapChart: software for the graphical presentation of linkage maps and QTLs. *J Hered* 2002, 93(1):77-78.

**Figures**
Figure 1

The genetic map and stable QTL for key fiber traits in the SXY No.1 x CSLZ RIL population

Figure 2

Collinearity between the SXY No.1 x CSLZ genetic map and physical map

Supplementary Files

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

- SupplementS1.pdf
- Fig.S1.png
- TableS8.xlsx
- TableS7.xlsx
• TableS6.xlsx
• TableS5.xlsx
• Additionalfiles.docx
• TableS4.xlsx
• TableS3.xlsx
• TableS2.xlsx
• TableS1.xlsx