

RESEARCH ARTICLE

QTL Mapping for Fiber and Yield Traits in Upland Cotton under Multiple Environments

Hantao Wang, Cong Huang, Huanle Guo, Ximei Li, Wenxia Zhao, Baosheng Dai, Zhenhua Yan, Zhongxu Lin*

National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University. Wuhan 430070, Hubei, China

* linzhongxu@mail.hzau.edu.cn



A population of 178 recombinant inbred lines (RILs) was developed using a single seed descendant from a cross between G. hirsutum. acc DH962 and G. hirsutum. cv Jimian5, was used to construct a genetic map and to map QTL for fiber and yield traits. A total of 644 polymorphic loci were used to construct a final genetic map, containing 616 loci and spanning 2016.44 cM, with an average of 3.27 cM between adjacent markers. Statistical analysis revealed that segregation distortion in the intraspecific population was more serious than that in the interspecific population. The RIL population and the two parents were phenotyped under 8 environments (two locations, six years), revealing a total of 134 QTL, including 64 for fiber qualities and 70 for yield components, independently detected in seven environments, explaining 4.40–15.28% of phenotypic variation (PV). Among the 134 QTL, 9 common QTL were detected in more than one environment, and 22 QTL and 19 new QTL were detected in combined analysis (E9). A total of 26 QTL hotspot regions were observed on 13 chromosomes and 2 larger linkage groups, and some QTL clusters related to fiber qualities or yield components were also observed. The results obtained in the present study suggested that to map accurate QTL in crops with larger plant types, such as cotton, phenotyping under multiple environments is necessary to effectively apply the obtained results in molecular marker-assisted selection breeding and QTL cloning.



OPEN ACCESS

Citation: Wang H, Huang C, Guo H, Li X, Zhao W, Dai B, et al. (2015) QTL Mapping for Fiber and Yield Traits in Upland Cotton under Multiple Environments. PLoS ONE 10(6): e0130742. doi:10.1371/journal.pone.0130742

Academic Editor: Kunbo Wang, Institute of Cotton Research of Chinese Academy of Agricultural Sciences, CHINA

Received: March 30, 2015
Accepted: May 24, 2015
Published: June 25, 2015

Copyright: © 2015 Wang et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: The authors have no support or funding to report.

Competing Interests: The authors have declared that no competing interests exist.

Introduction

Cotton (*Gossypium* L.) is an important economic crop, providing most of the natural textile fiber utilized worldwide. Upland cotton (*G. hirsutum*) is widely cultivated, and planted in more than 70 countries, contributing to over 95% of the total cotton yield worldwide [1]. Because of the softness and comfort of cotton fiber, cotton products are very popular. In recent decades, improvements in cotton fiber quality and yield have been stagnant and unable to meet the demands of the modern textile industry. However, yield is often negatively correlated with fiber quality in upland cotton [2,3]. Conventional cultivar breeding programs, primarily selecting novel allele combinations based on phenotypic selection [4], is difficult to break the linkage of negatively correlated traits. Fortunately, the development of genetic linkage maps facilitate



the dissection of quantitative trait loci (QTL) that control fiber qualities and yield components, which make it possible to pyramid elite genes of fiber quality and yield traits.

Because of the low genetic polymorphism between the intraspecific hybridization of upland cotton, several high-density interspecific linkage maps between *G. hirsutum* and *G. barbadense* have been constructed to study QTL for fiber quality and yield traits [5–7]. However, due to the sterility and segregation distortion of interspecific progeny, intraspecific hybridization has become the primary method in breeding programs, contributing to the recent development of upland cotton intraspecific genetic maps [8–15].

As the most important cotton cultivar, much research attention has been paid to the improvement of the fiber quality and yield of upland cotton. The QTL mapping of fiber quality traits could provide a solid foundation for future studies concerning marker-assisted selection breeding and map-based cloning. Hundreds of QTL associated with fiber quality and yield components have been obtained from F₂, F_{2:3}, and RIL populations in upland cotton [8,9,12– 25]. The F_2 population is a common population used in genetic map construction and QTL mapping, but these studies cannot be replicated; thus, stable and available QTL could not be identified using the F₂ population. The construction of immortalized mapping populations is an effective approach to obtain stable QTL, such as recombination inbred lines (RILs). However, the complex allotetraploid genome and agronomic traits in crops are inherited in a complex manner, suggesting that cotton traits are highly affected by environmental and climatic conditions, obtaining stable QTL in allotetraploid cotton is difficult. Tang et al. (2015) constructed a genetic map, containing 1,540 loci spanning 2,842.06 cM, and a total of 62 QTL were identified using combined analysis and single environment analysis; seventeen QTL were detected in more than one environment. Ning et al. (2014) identified 86 QTL for yield components and fiber qualities from an RIL population. In addition, a stable fiber strength QTL (qFS-D3-1), explaining 4.51–17.55% of the phenotypic variation (PV), and a stable fiber length QTL (qFL-D11-1), explaining 10.02-25.34% of the PV, were obtained.

In the present study, two upland cottons, DH962 and Jimian5, with different fiber qualities and yield component traits [21], were used as parents to establish a recombinant inbred line (RIL) population. The objectives of this study were to construct an intraspecific upland cotton map using SSRs, InDels and SNPs based on this RIL population, which was used to detect QTL associated with fiber quality and yield traits under multiple environmental conditions.

Materials and Methods

Mapping population and DNA isolation

The *G. hirsutum* acc. DH962 and *G. hirsutum* cv. Jimian5 were used as the mapping parents. DH962 was derived from the [(Jinmian6 × *G. thurberi*) F_4 × Jinmian6] F_3 population, showing good performance in fiber quality as a female parent and continuous self-pollination since 2001. Jimian5 is a cultivar with high yield as a male parent. DH962 and Jimian5 were crossed to obtain F_1 plants on the farm at Huazhong Agricultural University (HAU), Wuhan, China, in the summer of 2002 [10]. The F_1 plants were planted during the winter in Hainan Province and self-pollinated to produce the F_2 generation. The F_2 seeds were planted and self-pollinated to produce $F_{2:3}$ seeds on the farm of HAU in 2003. An RIL population was developed using the single seed descendant method to generate $F_{2:7}$ plants, which were subsequently planted at HAU for propagation in 2007. The $F_{7:8}$ generations of 178 RIL families were used in the present study. Genomic DNA was isolated from the fresh leaves of 178 RIL plants and parents using a CTAB procedure [26].



Fiber quality trait collection

The 178 RILs and parents were planted on a farm provided by Prof. Qizhong Xia at Huanggang Normal College, Huanggang (30.45° N, 114.93° E), Hubei, China in 2008 (E1), 2011 (E5), 2012 (E7) and 2013 (E8), and on a farm provided by Associate Prof. Dingguo Li at Yangtze University, Jingzhou (30.36° N, 112.15° E), Hubei, China in 2008 (E2), 2009 (E3), 2010 (E4) and 2011 (E6). These fields are only used for research purposes, and the field studies did not involve endangered or protected species. The fiber quality data in E7 and yield components in E1 were lost, and combined analysis (E9) was conducted after determining the mean values in seven environments. The lines were planted in single-row plots of 5 m in length with 0.8 m row spacing. All the lines were planted in the field using a randomized block. Twenty bolls from each line were simultaneously harvested for fiber quality and yield investigation. Six yield components and 5 fiber quality trait components were analyzed, including boll number per plant (BN), seed cotton weight per boll (SCW), lint weight per boll (LW), lint percentage (LP), lint index (LI), seed index (SI), fiber length (FL, mm), fiber strength (FS, cN/tex), fiber length uniformity ratio (FU), fiber elongation (FE), and micronaire (MIC).

DNA marker analysis

A total of 634 SSRs, InDels, and SNPs, selected according to Wang et al. (2015) [15], were used to genotype each RIL plant. PCR amplification and silver staining were performed according to Lin et al. (2005) [27]. The SRAP markers were not genotyped in this RIL population because these polymorphisms are difficult to crosstalk with those identified in previous studies. The PCR products of SSRs were separated on 6% denaturing polyacrylamide gels [27] or 8% native polyacrylamide gels using single-strand conformation polymorphism (SSCP) technology [28]. The PCR products of InDels and SNPs were separated on 8% native polyacrylamide gels using SSCP technology [28].

Data analysis, genetic map construction and QTL analysis

The difference between the two parents for each trait was detected using a t-test. The broadsense heritabilities of measured traits were calculated according to the method of Knapp et al. (1985) [29]. The coefficients of genetic correlation between measured traits were computed according to the method of Kwon and Torrie (1964) [30]. The phenotype data were analyzed using SPSS version 21.0 (SPSS, Chicago, IL, USA).

A chi-square test was performed to determine whether the genotypic frequencies at each locus deviated from the expected 1:1 segregation ratio in the RIL population. The genetic linkage map of the RIL population was constructed using JoinMap 3.0 [31] with a logarithm of odds (LOD) threshold of 4.0 and a maximum recombination fraction of 0.4. Map distances in centi-Morgans (cM) were calculated using the Kosambi mapping function [32]. Linkage groups were assigned to chromosomes based on BC₁ [33] and F₂ [10] linkage maps and marker mapping information on the CottonGen database (http://www.cottongen.org/). The linkage groups were named as "ChrX" based on the linkage group length from long to short. QTL were identified using Windows QTL Cartographer version 2.5 (http://statgen.ncsu.edu/qtlcart/WQTLCart.htm) based on composite interval mapping (CIM). The statistical significance of the LOD threshold value was determined using a permutation procedure (1,000 times for all traits). The QTL nomenclature was adapted according to the method of McCouch et al. (1997) [34]. The resulting linkage map and QTL were drawn using MapChart V2.2 software [35].



Table 1. The broad-sense heritabilities of fiber quality and yield component traits in the RIL population.

Components of variation	FL	FU	MIC	FE	FS	SCW	LW	LP	SI	BN	LI
σ^2	0.65	0.08	0.05	0.01	0.49	0.04	0.01	1.44	0.90	1.70	0.06
σ_{ge}^2	1.31	0.37	0.09	0.10	2.40	0.06	0.01	3.10	0.46	9.00	0.20
σ_{e}^{2}	1.15	1.13	0.12	0.21	2.80	0.23	0.04	6.60	0.38	70.00	0.53
H ² _B (%)	70.82	35.96	70.41	18.77	47.22	62.37	67.59	61.21	58.21	16.19	19.13

 σ^2 , the genotypic variance; σ^2_{qe} , the genotype and environment interaction variance; σ^2_{e} , the error variance; H^2_{e} , the broad-sense heritability

doi:10.1371/journal.pone.0130742.t001

Meta-analysis of the co-located QTL

The meta-analysis of the co-located QTL was conducted using Biomercator V3 software (http://moulon.inra.fr/index.php/fr/component/docman/cat_view/21-logiciels/101-abi-project-and-software/104-biomercator-v21). A map file and a QTL file were required to import into the Biomercator V3 software [36]. The map file contains map name, marker name, and distance between adjacent makers, etc. And the QTL file contains map name, QTL name, chromosome name, trait, LOD score, phenotypic variance (PV), position of the QTL, etc. The order 'meta-analysis' was used to integrate the QTL to detect the hotspot region.

Results

Trait performance and correlation analysis in the RIL population

The traits of fiber qualities and yield components are summarized in S1 Table. Overall, the trait values of DH962 were higher than those of Jimian5 in fiber qualities, and Jimian5 demonstrated higher trait values than DH962 in yield components, except that no significant difference was observed for SI and LI. The RIL population performed transgressive segregation on all traits. The results of ANOVA, shown in S2 Table revealed that most of fiber quality and yield component traits presented significant genetic and environmental effects (P < 0.01), except that SI showed no significantly environmental effect. And the broad-sense heritabilities of the fiber quality and yield component traits were showed in Table 1. The genetic potential of fiber quality and yield component traits were general low in cotton [37–39]. Boll number had the lowest broad-sense heritability (16.19%), fiber length had the highest broad-sense heritability (70.82%).

Genetic correlation analysis between fiber quality and yield component traits was calculated based on covariance (Table 2). FL was significantly and positively correlated with FS, FU, and significantly and negatively correlated with MIC, FE, LW, LP and BN. FS was significantly and positively correlated with FU and SI, and significantly and negatively correlated with MIC, FE, LW, LP, BN and LI. Among the yield component traits, most traits were positively correlated between two traits, except that SI was significantly and negatively correlated with LP.

Genetic map construction

A total of 644 loci were obtained from the selected 634 primer pairs from Wang et al. (2015) after genotyping the RIL population. After linkage analysis, 616 of 644 loci were mapped on 59 linkage groups; the total length of the linkage map was 2016.44 cM, with a mean distance of 3.27 cM between adjacent markers (Fig 1). The map included 538 SSR loci, 32 InDel loci and 46 SNP loci. Among the 59 linkage groups, there were 2–58 loci on each linkage group with 1.88–104.57 cM long. Fifty-three linkage groups were assigned to 24 chromosomes and 4 ones were unanchored; and most of the loci from two larger linkage groups (LG1-Chr1/15 and



Table 2. Genetic correlation analysis between fiber quality and yield component traits in the RIL popular

Trait	FL	FU	MIC	FE	FS	scw	LW	LP	SI	BN
FU	0.68**									
MIC	-0.71**	-0.12								
FE	-0.59**	-0.53**	0.14							
FS	0.99**	0.98**	-0.52**	-0.91**						
SCW	-0.08	0.01	0.20**	0.09	-0.05					
LW	-0.45**	-0.10	0.46**	0.28**	-0.43**	1.00**				
LP	-0.76**	-0.26**	0.62**	0.41**	-0.82**	0.17*	0.97**			
SI	0.12	0.05	-0.05	-0.18*	0.19*	0.15	0.03	-0.18*		
BN	-0.21**	0.08	0.87**	0.62**	-0.70**	-0.04	0.43**	0.95**	0.28**	
LI	-0.51**	-0.09	0.47**	0.49**	-0.46**	0.71**	0.99**	0.98**	-0.01	0.86**

^{*} Significance with P value of 0.05

doi:10.1371/journal.pone.0130742.t002

LG2-Chr9/23) were mapped on two pairs of homologous chromosomes (Chr1 and Chr15, Chr9 and Chr23, respectively).

Marker distorted segregation

All the 644 loci were tested for the expected Mendelian 1:1 segregation ratio, and 144 (22.36%) loci showed a distorted segregation ratio (P < 0.05). Among these, 73 loci favored the female parent 'DH962' allele, and 71 loci favored the male parent 'Jimian5' allele. A total of 134 loci showing segregation distortion were mapped, and the distorted loci were unevenly mapped on different chromosomes. Thirteen segregation distortion regions (SDRs) were identified on 9 chromosomes and 1 linkage group (LG1-Chr1/15, Chr3, Chr7, Chr9, Chr12, Chr17, Chr19, Chr20, Chr21, and Chr26) (Fig 1). Most of the distorted loci in a given SDR skewed towards the same parent allele; for example, two large SDRs were located on LG1-Chr1/15 and Chr26, which showed significant distortion towards 'DH962' allele (Fig 1).

QTL for fiber quality and yield component traits

A total of 134 QTL were detected on 21 chromosomes and 4 linkage groups, explaining 4.40–15.28% of the phenotypic variation (PV), with LOD scores ranging from 2.50 to 6.66 (S3 Table; Fig 1). In seven environments, 64 QTL for five fiber quality traits and 70 QTL for six yield components were identified. Among these QTL, 9 common QTL were detected in more than one environment. Among these 9 QTL, 5 QTL were associated with fiber qualities, and 4 QTL were associated with yield components. In addition, 22 of the 134 QTL and 19 new QTL (2 for FL, 1 for FS, 2 for MIC, 1 for FE, 3 for FU, 1 for SCW, 3 for LW, 2 for LP, 3 for BN, and 1 for LI) were identified in combined analysis (E9) (S3 Table).

Fiber length. Ten QTL were detected on 6 chromosomes (Chr4, Chr8, Chr10, Chr11, Chr17, and Chr21), with LOD scores ranging from 2.60 to 5.30, explaining 5.48–11.07% of the PV (S3 Table). Nine QTL derived from 'DH962' showed positive additive effects, and one QTL derived from 'Jimian5' showed a negative additive effect. The QTL qFL-c10-1 was detected in three individual environments (E5, E6 and E8), and this QTL, which was located between marker CIR305 and HAU-J5638, was also detected in combined analysis (E9), explaining 5.97–11.07% of the PV. The QTL qFL-c21-2 was detected in two environments (E5 and E8) and in

^{**} Significance with P value of 0.01



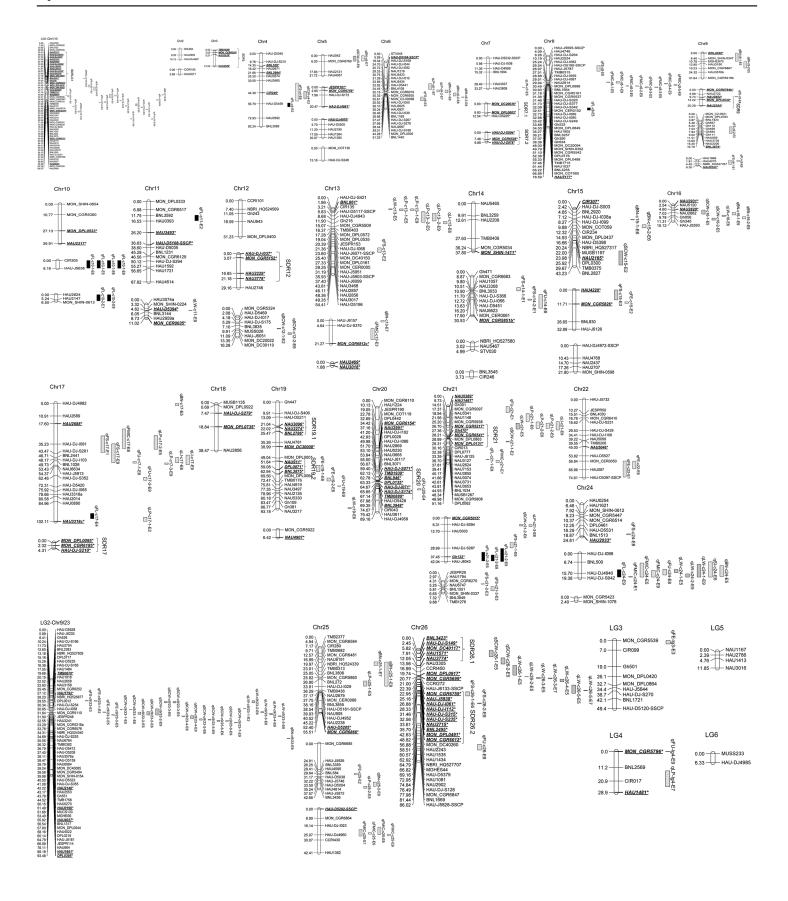




Fig 1. Genetic map and QTL for fiber quality and yield component traits in upland cotton. The map and QTL were detected under multiple environmental conditions in the RIL population derived from *G. hirsutum*. acc DH962 and *G. hirsutum*. cv Jimian5. The QTL were shown on the right of the Chromosomes/LGs. Markers showing segregation distortion are indicated as asterisks (*P < 0.05) and shown in underlined, italic, and bold text. The SDRs are shown in brackets.

doi:10.1371/journal.pone.0130742.g001

combined analysis (E9), and the QTL qFL-c10-2 was detected in E1 and combined analysis (E9).

Fiber strength. A total of 18 QTL for fiber strength were identified on 11 chromosomes and 2 linkage groups (LG1-Chr1/15, Chr4, Chr6, Chr8, Chr9, LG2-Chr9/23, Chr13, Chr14, Chr15, Chr17, Chr21, Chr24, and Chr26), explaining 4.86–11.39% of the PV, with LOD scores ranging from 2.52 to 5.18 (S3 Table). Among these, 11 QTL showed positive additive effects originating from 'DH962', and 7 QTL showed negative effects originating from 'Jimian5'. qFS-c17 was identified in E1 and combined analysis (E9), explaining 9.82 and 10.19% of the PV, with the LOD scores of 4.93 and 4.32, respectively.

Micronaire. Thirteen QTL were detected on 6 chromosomes and 1 linkage group (LG1-Chr1/15, Chr8, Chr13, Chr17, Chr21, Chr24, and Chr25), explaining 5.22–9.64% of the PV, with LOD scores ranging from 2.54 to 3.99 (S3 Table). The QTL qMIC-c24, originating from 'Jimian5', was identified in E1, E5 and combined analysis (E9), showing a negative additive effect and explaining 5.38% and 7.36% of the PV in E1 and E5, respectively. The QTL qMIC-c25 had a positive additive effect on MIC and was also detected in E1, E6, E8 and the combined analysis (E9). This QTL was located between the HAU-DJ-I023 and CCRI430 markers, explaining 5.80–7.20% of the PV. The QTL qMIC-c8-1, qMIC-c8-2 and qMIC-c8-3 were detected in both E5 and combined analysis (E9).

Fiber elongation. A total of 16 QTL were detected on 5 chromosomes and 3 linkage groups (LG1-Chr1/15, LG2-Chr9/23, Chr9, Chr15, Chr19, Chr22, Chr25, and LG3) (<u>S3 Table</u>). Among the 16 QTL, 9 QTL originated from 'DH962', and 7 QTL originated from 'Jimian5'. The QTL qFE-c1/15-3 was identified in E4 and combined analysis (E9), and the QTL qFE-c1/15-4 was identified in E5 and combined analysis (E9). The QTL qFE-c22 was observed in E1 and E6, showing a negative additive effect on E1 and a positive additive effect on E6, explaining 11.92% and 7.79% of the PV with LOD scores of 4.96 and 3.10, respectively.

Fiber length uniformity ratio. Seven QTL were detected on 7 chromosomes (Chr9, Chr14, Chr17, Chr20, Chr21, Chr24, and Chr26), explaining 5.09-8.67% of the PV, with LOD scores ranging from 2.52 to 4.13 (S3 Table). Only one QTL was located in the A_T genome, and six QTL were located in the D_T genome. Among which, 5 QTL showed positive additive effects originating from 'DH962', and 2 QTL showed negative effects originating from 'Jimian5'.

Seed cotton weight per boll. Nineteen QTL for SCW were detected on 7 chromosomes and 1 linkage group (Chr4, Chr5, LG2-Chr9/23, Chr12, Chr15, Chr16, Chr21, and Chr26), explaining 5.44–15.28% of the PV, with LOD scores ranging from 2.51 to 5.34 (S3 Table). Six QTL originated from the parent 'DH962', and 13 QTL originated from the parent 'Jimian5'. The QTL qSCW-c9/23-1 and qSCW-c9/23-2 were observed in E5 and E6, and the QTL qSCW-c9/23-2 was also detected in combined analysis (E9). The QTL qSCW-c9/23-1 explained 5.57 and 5.74% of the PV, with LOD scores of 2.51 and 2.80, respectively. The QTL qSCW-c9/23-2 explained 5.92–7.82% of the PV, with LOD scores ranging from 3.20 to 3.99. Four QTL, qSCW-c9/23-3, qSCW-c9/23-4, qSCW-c9/23-5, and qSCW-c9/23-6, were detected in both E6 and combined analysis (E9).

Lint weight per boll. For lint weight per boll, 20 QTL were detected on 7 chromosomes and 2 linkage groups (LG1-Chr1/15, Chr7, LG2-Chr9/23, Chr13, Chr16, Chr19, Chr21, Chr24, and Chr26), explaining 4.40–9.93% of the PV, with LOD scores ranging from 2.50 to 4.70 (<u>S3 Table</u>). Ten QTL originating from 'DH962' showed positive additive effects on LW, and the



other ten QTL originating from 'Jimian5' showed negative additive effects on LW. Two QTL, qLW-c26-3 and qLW-c26-5, were detected in both E5 and E7, both originating from the 'DH962' parent. The QTL qLW-c26-3 explained 9.36 and 9.48% of the PV, with LOD scores of 3.17 and 3.15, respectively. The QTL qLW-c26-5 explained 9.93 and 5.54% of the PV, with LOD scores of 3.58 and 2.53, respectively. qLW-c24-1 and qLW-c24-2 were identified in combined analysis (E9) as one QTL.

Lint percentage. A total of 13 QTL for LP were identified on 5 chromosomes and 2 linkage groups (LG1-Chr1/15, Chr6, Chr13, Chr17, Chr21, Chr25, and LG4), explaining 4.48–12.50% of the PV, with LOD scores range from 2.52 to 6.66 (S3 Table). Five QTL derived from 'DH962' showed positive additive effects, whereas eight QTL derived from 'Jimian5' showed negative additive effects. Two QTL, qLP-c13-1 and qLP-c13-2, were detected in both E5 and combined analysis (E9). The QTL qLP-c21-2 was detected in E6 and combined analysis (E9).

Boll number per plant. Fourteen QTL for BN were detected on 8 chromosomes and 2 linkage groups (LG1-Chr1/15, Chr5, Chr6, Chr9, LG2-Chr9/23, Chr13, Chr15, Chr17, Chr24, and Chr25), with LOD scores ranging from 2.52 to 4.07, explaining 5.15–9.08% of the PV (S3 Table). Eight QTL derived from 'DH962'showed positive additive effects, whereas six QTL derived from 'Jimian5' showed negative additive effects. The QTL qBN-c25-1 was detected in E7 and combined analysis (E9).

Seed index and lint index. Two and two QTL were identified for SI and LI, respectively (S3 Table). Two QTL of SI explained 6.90 and 14.38% of the PV, with LOD scores of 3.49 and 6.64, respectively. Two QTL for LI explained 5.02 and 7.30% of the PV, with LOD scores of 2.69 and 3.71, respectively. The QTL qLI-c17 was detected in E6 and combined analysis (E9).

QTL clusters

A total of 24 co-located QTL regions were observed on 13 chromosomes and 2 larger linkage groups (Chr4, Chr5, Chr6, Chr9, Chr10, Chr13, Chr14, Chr17, Chr21, Chr22, Chr24, Chr25, Chr26, LG1-Chr1/15, and LG2-Chr9/23). After QTL meta-analysis, 26 QTL hotspot regions were obtained (S4 Table; S1 Fig). Except for 9 common QTL, 17 QTL hotspot regions affected different fiber quality or yield component traits (S1 Fig). For example, on LG2-Chr9/23, there were 3 QTL related to different traits (FS, FE, and LW) distributed on a region of 0.46 cM (S4 Table; S1 Fig). Several QTL clusters were distributed on 15 chromosomes and 2 linkage groups (Chr4, Chr5, Chr6, Chr8, Chr9, Chr10, Chr13, Chr14, Chr16, Chr17, Chr21, Chr22, Chr24, Chr25, Chr26, LG1-Chr1/15, and LG2-Chr9/23). On LG2-Chr9/23, 18 QTL included 4 traits (FS, FE, SCW, and LW) were located on a region ranging from 20.1 to 40.8 cM. Among the 18 QTL, 8 QTL were associated with SCW, 5 QTL were associated with LW, 3 QTL were associated with FE, and 1 QTL was associated with FS.

Discussion

In the present study, 616 loci were mapped on 59 linkage groups, and the total length of the linkage map was 2016.44 cM, covering 40.33% of the upland cotton genome. Compared with high-density interspecific genetic maps [40–43], the intraspecific genetic map generated in the present study showed low coverage throughout the cotton genome. This phenomenon was also observed for the construction of other genetic maps in upland cotton [11,13,14]. Due to the narrow genetic base of upland cotton, we screened most of the SSRs in the CottonGen database (http://www.cottongen.org) and some SNPs and InDels developed in the laboratory [15], but the results were not satisfactory, and with the increase in map density, the efficiency of QTL detection was greatly improved [7,15]. There is no high-density and coverage genetic map for upland cotton, which is the huge obstacle for QTL mapping in upland cotton. In previous



studies [14,15], the effect of SSRs was lower in the construction of the upland cotton genetic map, and the development of lots of SNPs and InDels using next-generation sequencing technology could provide a new outlet for future studies.

Segregation distortion widely exists in the study of population genetics. In the present study, 144 (22.36%) of the 644 loci showed segregation distortion. In previous studies of cotton [3,8,11,13,14,19,27,42,44-50], many reports have focused on segregation distortion, we summarized these results in <u>\$5 Table</u>. The percentages of distorted loci in interspecific populations ranged from 7.65% to 25.50%, while they ranged from 22.36% to 81.25% in intraspecific populations. Although the genetic difference is bigger, the segregation distortion is smaller in interspecific populations, and this phenomenon presents an interesting topic for future studies. §5 Table also showed that most of the distorted loci skewed toward upland cotton in interspecific populations, and most of the distorted loci skewed toward one of the parents in upland cotton in intraspecific populations. Lacape et al. (2010) showed that the average fiber characteristics of the interspecific RILs were closer to the G. hirsutum parent than to the G. barbadense parent, and 71% and 29% of the observed distorted allelic skewed toward G. hirsutum and G. barbadense, respectively. As the important cultivated cotton, most of the varieties of upland cotton were improved, and some good genes from other Gossypium species were introgressed into upland cotton through hybridization and backcross methods. The introgression of good genes generated the genetic differences observed between the mapping parents of upland cotton, resulting in segregation distortion [8,11,13,14,19]. The number of distorted loci skewed toward 'DH962' and 'Jimian5' were similar, and the percentage of distorted loci was lower than that in other upland cotton populations in the present study, suggesting that the ability of selection and the combination of gametophyte mapping parents are similar. In addition, the fiber quality and yield component traits were significantly different (S1 Table), and all the results showed that the two parents were suitable for the development of the RIL population.

The clustering of QTL in tetraploid cotton has been reported in some studies [3,8,11,13,14,19,20,47,51,52]. The present study also identified 26 QTL hotspot regions, among which 17 QTL hotspot regions affected two or more different fiber quality or yield component traits. The phenomenon of QTL clustering might represent the linkage of genes and QTL or result from pleiotropic effects of a single QTL in the same genomic region [3]. For example, NAU5107, detected on E5, was the nearest marker of qFS-c1/15-1 and qFE-c1/15-4; and FS was significantly and negatively correlated with FE in the present study (Table 2). These QTL hotspot regions revealed that the linkage drag of QTL hindered the improvement of fiber quality [20]. In addition, 11 larger QTL clusters were also obtained in the present study. As shown in Fig 1, we observed that most of the clusters showed the enrichment of fiber quality or yield component traits. On LG1-Chr1/15, most of the QTL were correlated with FS and FE in the cluster. On Chr13, most of the QTL were correlated with LP and BN in the cluster. On LG2-Chr9/23, 18 QTL were distributed on the QTL clusters, 8 QTL were associated with SCW, and 5 QTL were associated with LW. Previous studies [51,52] generated multiple QTL clusters and hotspots of fiber quality or yield component traits in the cotton genome through the analysis of most of the publications on cotton. The QTL clusters provided valuable information to determine genome regions with different traits.

<u>S2 Table</u> showed that all the fiber quality and yield component traits presented significant environmental effects, and the present study revealed that cotton traits were highly affected by environments and climate conditions. An example of meteorological information of E5, E6, E7 and E8 is shown in <u>S6 Table</u> and <u>S2 Fig. S6 Table</u> showed that the number of rainy day during the four years were different. For example, 14 rainy days were reported in E7, but 6 rainy days were reported in E8 in August. And <u>S7 Table</u> showed that the temperature presented significant environmental effects (P < 0.01) in July and August, and these two months represented a



critical period for blossom and fiber development in cotton. The different climate conditions during the two months might have seriously affected the fiber quality and yield of cotton.

Because the genetic maps were constructed based on different populations and genetic backgrounds, and environmental elements would affect the expression of QTL, it was difficult to obtain common QTL among different populations. However, lots of QTL observed in the present study were still identified according to the same markers of the same chromosomes in other reports. The QTL qLP-c6-2 was detected as QTL-lp2 in the F₂ population from a cross between TM-1 and T586 [53]. QTL qFS-c14-1 corresponded to the QTL qFS-D2-1 in the 4WC population [20]. QTL qFE-c1/15-2, qFE-c9/23-1, qFL-c17, qFS-c9/23, and qFU-c17-1 were detected as qFE-A1-1, qFE-A9-1, qFL-D3-1, qFS-A9-1, and qFU-D3-1, respectively, in the RIL population resulting from a cross between Prema and 86-1 [13]. The QTL qFL-c10-2 was detected as the major QTL qFL10.1 in the RIL population resulting from a cross between the upland cotton cultivars Yumian 1 and 7235 [14]. QTL qFL-c10-2 was also detected as a major QTL qFL10.1 in the report of Shao et al (2014); and qFL-c8, qFMIC-c1/15-1, qFMIC-c8-5, qFE-c25, qFS-c9-2, and qFS-c21-3 were detected as qFL08.1, qFM01.1, qFM08.1, qFE25.1, qFS09.1, and qFS21.2 in this research [25], respectively. The QTL qFMIC-c1/15-2 was detected as qMC-07A-c1-1 in an interspecific population [3]. And qFE-c15 was the same as the QTL qFE15.2 in a three-parent composite population of upland cotton [12]. QTL qFE-lg3 and qFEc22 were detected as qELO-c18-2 and qELO-c22 in a random-mated recombinant inbred population [54], respectively. At the same time, qBN-c6-1 was detected as qNB-A6-1 by linkage analysis and association mapping [55]. The QTL qLP-c13-1 and qMIC-c24 were detected as qLP-c13 and qMV-c24, respectively, in a F₂ population based on the same parents used in the present study [15]. These QTL across different populations could be examined to obtain candidate genes of related traits, and the adjacent markers of these QTL could be used for markerassisted selection. Some common makers (BNL2495, BNL4028, BNL4108, Gh584, HAU2147, MON_CGR5826, MON_CGR6019, MON_CGR6764, MON_SHIN0613, NAU2820, NAU3774, NAU3308, NAU3948, NAU4891, and NAU5433) were also identified in other cotton studies [3,9,12-15,20,48,53,56]. For example, BNL4108 was associated with BN and LP in the present study, but associated with FL, FU, FE, and LP in previous studies [9,53]. NAU3308 was associated with FS in the present study, but associated with FL, FS, SI, and LP in a previous study [20]. MON CGR5826 was associated with FE and FS, but associated with FU in the report of Zhang et al. (2012) [12]. The common markers associated with different traits might result in the development of QTL clusters and linkage drag.

Among the 134 QTL detected in the present study, 9 common QTL were obtained in more than one environment. The difficulty of obtaining stable QTL has been reported in the previous studies [8,11,12,14]. Two QTL, qFL-c10-1 and qMIC-c25, were detected in three environments and combined analysis (E9), and the results revealed that fiber quality traits were more stable than yield traits under multiple environmental conditions. Notably, two stable QTL were not clustered with any other QTL associated with fiber quality and yield traits, providing an opportunity to identify candidate genes for improving the fiber quality of upland cotton.

Supporting Information

S1 Fig. QTL Meta-analysis of hotspot regions using BiomercatorV3 software. (TIF)

S2 Fig. Average temperatures in July, August and September in E5, E6, E7 and E8. (TIF)



S1 Table. Phenotypic value of fiber quality and yield component traits in the RIL population and the parents.

(XLS)

S2 Table. Variance analysis of fiber quality and yield component traits in the mapping population.

(XLS)

S3 Table. QTL for fiber quality and yield component traits in the RIL population under multiple environmental conditions.

(XLS)

S4 Table. The AIC scores of the QTL hotspot regions analyzed by Biomercator V3 software. (XLS)

S5 Table. Characteristics of distorted loci in interspecific and intraspecific populations. (XLS)

S6 Table. Weather characteristics in July, August and September in E5, E6, E7 and E8. (XLS)

S7 Table. Variance analysis of the average temperatures in July, August and September in E5, E6, E7 and E8.

(XLS)

Author Contributions

Conceived and designed the experiments: ZL. Performed the experiments: HW CH HG XL WZ BD ZY. Analyzed the data: HW. Wrote the paper: HW ZL.

References

- Chen ZJ, Scheffler BE, Dennis E, Triplett BA, Zhang TZ, Guo WZ, et al. Toward sequencing cotton (Gossypium) genomes. Plant Physiol. 2007; 145(4): 1303–1310. PMID: 18056866
- Meredith W. Registration of MD 52ne high fiber quality cotton germplasm and recurrent parent MD 90ne. Crop Sci. 2005; 45(2): 807–808.
- Yu JW, Zhang K, Li SY, Yu S, Zhai HH, Wu M, et al. Mapping quantitative trait loci for lint yield and fiber quality across environments in a Gossypium hirsutum x Gossypium barbadense backcross inbred line population. Theor Appl Genet. 2013; 126(1): 275–287. doi: 10.1007/s00122-012-1980-x PMID: 23064252
- Bridge RR, Meredith W. Comparative performance of obsolete and current cotton cultivars. Crop Sci. 1983; 23(5): 949–952.
- Li XM, Yuan DJ, Wang HT, Chen XM, Wang B, Lin ZX, et al. Increasing cotton genome coverage with polymorphic SSRs as revealed by SSCP. Genome. 2012; 55(6): 459–470. doi: 10.1139/g2012-032 PMID: 22670804
- **6.** Lacape JM, Gawrysiak G, Cao TV, Viot C, Llewellyn D, Liu SM, et al. Mapping QTLs for traits related to phenology, morphology and yield components in an inter-specific *Gossypium hirsutum* × *G. barbadense* cotton RIL population. Field Crops Res. 2013; 144: 256–267.
- Yu JZ, Ulloa M, Hoffman SM, Kohel RJ, Pepper AE, Fang DD, et al. Mapping genomic loci for cotton plant architecture, yield components, and fiber properties in an interspecific (Gossypium hirsutum L. x G. barbadense L.) RIL population. Mol Genet Genomics. 2014; 289(6): 1347–1367. doi: 10.1007/s00438-014-0930-5 PMID: 25314923
- 8. Shen XL, Guo WZ, Zhu XF, Yuan YL, Yu JZ, Kohel RJ, et al. Molecular mapping of QTLs for fiber qualities in three diverse lines in Upland cotton using SSR markers. Mol Breeding. 2005; 15(2): 169–181.
- Zhang ZS, Xiao YH, Luo M, Li XB, Luo XY, Hou L, et al. 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.



- Lin ZX, Zhang YX, Zhang XL, Guo XP. A high-density integrative linkage map for Gossypium hirsutum. Euphytica. 2009; 166(1): 35–45.
- Zhang ZS, Hu MC, Zhang J, Liu DJ, Zheng J, Zhang K, et al. 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.
- Zhang K, Zhang J, Ma J, Tang SY, Liu DJ, Teng ZH, et al. Genetic mapping and quantitative trait locus analysis of fiber quality traits using a three-parent composite population in upland cotton (Gossypium hirsutum L.). Mol Breeding. 2012; 29(2): 335–348. doi: 10.1089/neu.2011.1862 PMID: 21806472
- Ning ZY, Chen H, Mei HX, Zhang TZ. Molecular tagging of QTLs for fiber quality and yield in the upland cotton cultivar Acala-Prema. Euphytica. 2014; 195(1): 143–156.
- 14. Tang SY, Teng ZH, Zhai TF, Fang XF, Liu F, Liu DJ, 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. doi: 10.1007/s00359-014-0964-5 PMID: 25398576
- Wang HT, Jin X, Zhang BB, Shen C, Lin ZX. Enrichment of an intraspecific genetic map of upland cotton by developing markers using parental RAD sequencing. DNA Research. 2015; 22(2):147–160. doi: 10.1093/dnares/dsu047 PMID: 25656006
- Zhang TZ, Yuan YL, Yu J, Guo WZ, Kohel RJ. Molecular tagging of a major QTL for fiber strength in Upland cotton and its marker-assisted selection. Theor Appl Genet. 2003; 106(2): 262–268. PMID: 12582851
- Shen XL, Zhang TZ, Guo WZ, Zhu XF, Zhang XY. Mapping fiber and yield QTLs with main, epistatic, and QTL x environment interaction effects in recombinant inbred lines of Upland cotton. Crop Sci. 2006: 46: 61.
- **18.** Wang BH, Guo WZ, Zhu XF, Wu YT, Huang NT, Zhang TZ. QTL mapping of fiber quality in an elite hybrid derived-RIL population of upland cotton. Euphytica. 2006; 152(3): 367–378.
- Shen XL, Guo WZ, Lu QX, Zhu XF, Yuan YL, Zhang TZ. Genetic mapping of quantitative trait loci for fiber quality and yield trait by RIL approach in Upland cotton. Euphytica. 2007; 155(3): 371–380.
- Qin HD, Guo WZ, Zhang YM, Zhang TZ. QTL mapping of yield and fiber traits based on a four-way cross population in Gossypium hirsutum L. Theor Appl Genet. 2008; 117(6): 883–894. doi: 10.1007/ s00122-008-0828-x PMID: 18604518
- Lin ZX, Feng CH, Guo XP, Zhang XL. Genetic analysis of major QTLs and epistasis interaction for yield and fiber quality in upland cotton. Sci Agr Sinica. 2009; 42(9): 3036–3047.
- Chen H, Qian N, Guo WZ, Song QP, Li BC, Deng FJ, et al. Using three selected overlapping RILs to fine-map the yield component QTL on Chro.D8 in Upland cotton. Euphytica. 2010; 176(3): 321–329.
- 23. Liu RZ, Wang BH, Guo WZ, Qin YS, Wang LG, Zhang YM, et al. Quantitative trait loci mapping for yield and its components by using two immortalized populations of a heterotic hybrid in Gossypium hirsutum L. Mol Breeding. 2012; 29(2): 297–311.
- Sun FD, Zhang JH, Wang SF, Gong WK, Shi YZ, Liu AY, et al. QTL mapping for fiber quality traits across multiple generations and environments in upland cotton. Mol Breeding. 2012; 30(1): 569–582.
- Shao QS, Zhang FJ, Tang SY, Liu Y, Fang XM, Liu DX, et al. Identifying QTL for fiber quality traits with three upland cotton (Gossypium hirsutum L.) populations. Euphytica. 2014; 198(1): 4358.
- Paterson AH, Brubaker CL, Wendel JF. A rapid method for extraction of cotton (Gossypium spp.) genomic DNA suitable for RFLP or PCR analysis. Pl Mol Biol Rep. 1993; 11(2): 122–127.
- 27. Lin ZX, He DH, Zhang XL, Nie YC, Guo XP, Feng CH, et al. Linkage map construction and mapping QTL for cotton fibre quality using SRAP, SSR and RAPD. Plant Breed. 2005; 124(2): 180–187.
- Wang XQ, Ren GF, Li XM, Tu JL, Lin ZX, Zhang XL. Development and evaluation of intron and insertion-deletion markers for Gossypium barbadense. PI Mol Biol Rep. 2012; 30(3): 605–613.
- Knapp SJ, Stroup WW, Ross WM. Exact confidence intervals for heritability on a progeny mean basis. Crop Sci. 1985; 25(1): 192–194.
- Kwon SH, Torrie JH. Heritability and interrelationship among traits of two soybean populations. Crop Sci. 1964; 4(2): 196–198.
- Stam P. Construction of integrated genetic linkage maps by means of a new computer package: Join Map. Plant J. 1993; 3(5): 739–744.
- **32.** Kosambi D. The estimation of map distances from recombination values. Ann Eugenics. 1943; 12(1): 172–175.
- Li XM. Construction of introgression lines, development of new markers, and construction of high-density genetic linkage map in cotton [Doctor dissertation]. Wuhan: Huazhong Agricultural University; 2013.



- 34. McCouch SR, Cho YG, Yano M, Paul E, Blinstrub M, Morishima H, et al. Report on QTL nomenclature. Rice Genet Newsl. 1997; 14(11): 11–131. PMID: 9433552
- Voorrips R. MapChart: software for the graphical presentation of linkage maps and QTLs. J Hered. 2002; 93(1): 77–78. PMID: 12011185
- Arcade A, Labourdette A, Falque M, Mangin B, Chardon F, Charcosset A, et al. BioMercator: integrating genetic maps and QTL towards discovery of candidate genes. Bioinformatics. 2004; 20(14): 2324– 2326. PMID: 15059820
- 37. Tang B, Jenkins JN, Watson CE, McCarty JC, Creech RG. Evaluation of genetic variances, heritabilities, and correlations for yield and fiber traits among cotton F₂ hybrid populations. Euphytica. 1996; 91: 315–322.
- **38.** McCarty JC, Jenkins JN, Wu JX. Primitive accession derived germplasm by cultivar crosses as sources for cotton improvement. Crop Sci. 2004; 44(4): 1226–1230.
- **39.** Naveed M, Azhar FM, Ali A. Estimates of heritabilities and correlations among seed cotton yield and its components in *Gossypium hirsutum* L. Int J Agri Biol. 2004; 6(4): 712–714.
- Rong JK, Abbey C, Bowers JE, Brubaker CL, Chang C, Chee PW, et al. A 3347-locus genetic recombination map of sequence-tagged sites reveals features of genome organization, transmission and evolution of cotton (Gossypium). Genetics. 2004; 166(1): 389–417. PMID: 15020432
- 41. Fang DD, Yu JZ. Addition of 455 microsatellite marker loci to the high-density Gossypium hirsutum TM-1 x G. barbadense 3–79 genetic map. J Cotton Sci. 2012; 16: 229–248.
- **42.** Zhao L, Lv YD, Cai CP, Tong XC, Chen XD, Zhang W, et al. Toward allotetraploid cotton genome assembly: integration of a high-density molecular genetic linkage map with DNA sequence information. BMC Genomics. 2012; 13(1): 539.
- **43.** Wang HT, Li XM, Gao WZ, Jin X, Zhang XL, Lin ZX. Comparison and development of EST–SSRs from two 454 sequencing libraries of *Gossypium barbadense*. Euphytica. 2014; 198(2): 277–288.
- 44. Lacape JM, Nguyen TB, Thibivilliers S, Bojinov B, Courtois B, Cantrell RG, et al. A combined RFLP-SSR-AFLP map of tetraploid cotton based on a Gossypium hirsutum x Gossypium barbadense backcross population. Genome. 2003; 46(4): 612–626. PMID: 12897870
- 45. Mei M, Syed NH, Gao W, Thaxton PM, Smith CW, Stelly DM, et al. Genetic mapping and QTL analysis of fiber-related traits in cotton (Gossypium). Theor Appl Genet. 2004; 108(2): 280–291. PMID: 14513220
- 46. Yu JW, Yu SX, Lu CR, Wang W, Fan SL, Song MZ, et al. High-density linkage map of cultivated allote-traploid cotton based on SSR, TRAP, SRAP and AFLP markers. J Integr Plant Biol. 2007; 49(5): 716–724.
- **47.** He DH, Lin ZX, Zhang XL, Nie YC, Guo XP, Zhang YX, 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* × *Gossypium barbadense*. Euphytica. 2007; 153(1–2): 181–197.
- **48.** Chen L, Zhang ZS, Hu MC, Wang WW, Zhang J, Liu DJ, et al. Genetic linkage map construction and QTL mapping for yield and fiber quality in upland cotton (*Gossypium hirsutum* L.). Acta Agron Sinica. 2008; 7: 015.
- 49. Zhang YX, Lin ZX, Xia QZ, Zhang MJ, Zhang XL. Characteristics and analysis of simple sequence repeats in the cotton genome based on a linkage map constructed from a BC₁ population between Gossypium hirsutum and G. barbadense. Genome. 2008; 51(7): 534–546. doi: 10.1139/G08-033 PMID: 18545277
- 50. Yu Y, Yuan DJ, Liang SG, Li XM, Wang XQ, Lin ZX, et al. Genome structure of cotton revealed by a genome-wide SSR genetic map constructed from a BC₁ population between gossypium hirsutum and G. barbadense. BMC Genomics. 2011; 12(1): 15.
- Said JI, Lin ZX, Zhang XL, Song MZ, Zhang JF. A comprehensive meta QTL analysis for fiber quality, yield, yield related and morphological traits, drought tolerance, and disease resistance in tetraploid cotton. BMC Genomics. 2013; 14(1): 776.
- Said JI, Song MZ, Wang HT, Lin ZX, Zhang XL, Fang DD, et al. A comparative meta-analysis of QTL between intraspecific Gossypium hirsutum and interspecific G. hirsutum x G. barbadense populations. Mol Genet Genomics. 2014: 1–23.
- Guo WZ, Ma GJ, Zhu YC, Yi CX, Zhang TZ. Molecular tagging and mapping of quantitative trait loci for lint percentage and morphological marker genes in upland cotton. J Integr Plant Biol. 2006; 48(3): 320–326.
- 54. Fang DD, Jenkins JN, Deng DD, McCarty JC, Li P, Wu JX. Quantitative trait loci analysis of fiber quality traits using a random-mated recombinant inbred population in Upland cotton (Gossypium hirsutum L.).
 BMC Genomics. 2014; 15: 397. doi: 10.1186/1471-2164-15-397
 PMID: 24886099



- 55. Zhang TZ, Qian N, Zhu XF, Chen H, Wang S, Mei HX, et al. Variations and transmission of QTL alleles for yield and fiber qualities in upland cotton cultivars developed in China. Plos One. 2013; 8: e57220. doi: 10.1371/journal.pone.0057220 PMID: 23468939
- 56. Wang XQ, Yu Y, Sang J, Wu QZ, Zhang XL, Lin ZX. Intraspecific linkage map construction and QTL mapping of yield and fiber quality of Gossypium babrdense. Aust J Crop Sci. 2013; 7: 1252.