

Citation: Zhai H, Gong W, Tan Y, Liu A, Song W, Li J, et al. (2016) Identification of Chromosome Segment Substitution Lines of *Gossypium barbadense* Introgressed in *G. hirsutum* and Quantitative Trait Locus Mapping for Fiber Quality and Yield Traits. PLoS ONE 11(9): e0159101. doi:10.1371/journal. pone.0159101

Editor: David D Fang, USDA-ARS Southern Regional Research Center, UNITED STATES

Received: December 27, 2015

Accepted: June 27, 2016

Published: September 7, 2016

Copyright: © 2016 Zhai et al. This is an open access article distributed under the terms of the <u>Creative</u> <u>Commons Attribution License</u>, 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: This work was funded by the National Natural Science Foundation of China (31101188) to YZS, the National High Technology Research and Development Program of China (2012AA101108) to YLY, and the National Basic Research Program of China (973 Project) (2010CB126000) to YLY.

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

RESEARCH ARTICLE

Identification of Chromosome Segment Substitution Lines of *Gossypium barbadense* Introgressed in *G. hirsutum* and Quantitative Trait Locus Mapping for Fiber Quality and Yield Traits

Huanchen Zhai^{1,2°}, Wankui Gong^{1°}, Yunna Tan^{1°}, Aiying Liu¹, Weiwu Song¹, Junwen Li¹, Zhuying Deng¹, Linglei Kong¹, Juwu Gong¹, Haihong Shang¹, Tingting Chen¹, Qun Ge¹, Yuzhen Shi¹*, Youlu Yuan¹*

1 State Key Laboratory of Cotton Biology, Key Laboratory of Biological and Genetic Breeding of Cotton, The Ministry of Agriculture, Institute of Cotton Research, Chinese Academy of Agricultural Sciences. Anyang, Henan, China, 2 College of Bioengineering, Henan University of technology, Zhengzhou, Henan, China

These authors contributed equally to this work.
* youluyuan@hotmail.com (YLY); shiyzmb@126.com (YZS)

Abstract

Chromosome segment substitution lines MBI9804, MBI9855, MBI9752, and MBI9134, which were obtained by advanced backcrossing and continuously inbreeding from an interspecific cross between CCRI36, a cultivar of upland cotton (Gossypium hirsutum) as the recurrent parent, and Hai1, a cultivar of sea island cotton (G. barbadense) as the donor parent, were used to construct a multiple parent population of (MBI9804×MBI9855)×(M-BI9752×MBI9134). The segregating generations of double-crossed F_1 and F_2 and $F_{2:3}$ were used to map the quantitative trait locus (QTL) for fiber quality and yield-related traits. The recovery rate of the recurrent parent CCRI36 in the four parental lines was from 94.3%-96.9%. Each of the parental lines harbored 12-20 introgressed segments from Hai1across 21 chromosomes. The number of introgressed segments ranged from 1 to 27 for the individuals in the three generations, mostly from 9 to 18, which represented a genetic length of between 126 cM and 246 cM. A total of 24 QTLs controlling fiber quality and 11 QTLs controlling yield traits were detected using the three segregating generations. These QTLs were distributed across 11 chromosomes and could collectively explain 1.78%-20.27% of the observed phenotypic variations. Sixteen QTLs were consistently detected in two or more generations, four of them were for fiber yield traits and 12 were for fiber guality traits. One introgressed segment could significantly reduce both lint percentage and fiber micronaire. This study provides useful information for gene cloning and marker-assisted breeding for excellent fiber quality.

Introduction

Cotton is one of the most important cash crops in the world and cotton fiber provides the main natural raw material for the textile industry. Upland cotton (*Gossypium hirsutum*) has a high yield and wide adaptability, but a relatively low fiber quality. On the other hand, sea island cotton (*G. barbadense*) has excellent fiber quality with low yield and limited adaptability. Therefore, one way to improve fiber quality of upland cotton is to introgress the favorable genes from sea island cotton to upland cotton. However, the yield and quality of cotton are quantitative traits that are affected by multiple genes and often negatively correlated [1-3]. Therefore, the simultaneous improvement of cotton fiber quality and yield is a tall task for breeders in conventional breeding [4]. With the continuous development and improvement of molecular marker technologies, researchers have conducted extensive studies to construct cotton genetic maps and identify quantitative trait locus (QTL). This would make it possible to simultaneously improve both of fiber quality and yield in a breeding program.

The construction of the first cotton molecular genetic map [5] had facilitated QTL mapping. Completion of the cotton genome draft sequence laid a foundation for further molecular design breeding at the whole genomic level [6–9]. Most segregating populations for QTL identification are F_2 [10–12], BC₁ [13], BIL [14] and RIL [15–18]. However, because these segregating populations (e.g., F_2) are usually not immortal, the results of QTL identification are usually difficult to repeat. Furthermore, although several QTLs for cotton fiber quality or yield traits have been identified, fine mapping and cloning of these genes have rarely begun. Chromosome segment substitution lines (CSSLs), also known as introgression lines, are permanent populations that possess the same genetic background as the recurrent parent. Differences among CSSLs usually involve only one or a few of the introgressed chromosome segments, which in turn effectively eliminate interference of the genetic background. CSSLs are also highly efficient in detecting QTLs with minor effects. Therefore, CSSLs are ideal materials for QTL fine mapping, gene cloning, and investigating QTL interactions. Since Eshed and Zamir first constructed introgression lines of tomato [19], these have been successfully applied in rice, corn, and other plants [20–23].

CSSLs are seldom reported in QTL studies in cotton. Stelly et al. first constructed 17 chromosome substitution lines of *G. barbadense* in TM-1 background of *G. hirsutum* [24]. Subsequently, the same research team performed a thorough analysis of the genetic effects of CSSLs [25–31]. Their results showed that the sea island cotton genotype has positive effects on fiber quality traits, suggesting that these particular traits are influenced by multiple genes [32]. Other researchers also used CSSLs for QTL mapping of fiber quality and yield traits [33–37]. Although these CSSL populations are beneficial for QTL mapping, a large gap between the QTL mapping and the application of these lines in breeding programs remains to be resolved.

To simultaneously obtain brand-new lines for direct application in breeding while conducting basic researches, we constructed a CSSL population using upland cotton CCRI36 and sea island cotton Hai1, both of which are commercially grown cultivars. A genetic linkage map containing 2,292 markers was constructed [38] and cotton fiber quality and yield-related QTLs were identified in this CSSL population. Liang et al. [39] detected 20 yield-related QTLs in BC_5F_2 of this CSSL population. Through multi-ecological environment evaluations of the yield and fiber quality of the population (BC_5F_3 , $BC_5F_{3:4}$, and $BC_5F_{3:5}$), Zhang et al. [40] and He et al. [41] mapped specific QTLs for fiber quality and yield-related traits using selected lines with stable and excellent fiber quality and yield.

In the present study, based on the phenotypic performance of this CSSL population [38], as well as previous findings from multi-ecological environment investigations [38–41], four introgression lines MBI9804, MBI9855, MBI9752, and MBI9134 with excellent fiber quality

were selected as parental lines, and a double-crossed population of (MBI9804×MBI9855)× (MBI9752×MBI9134) was constructed. The introgressed Hai1 segments were evaluated in the segregating generations F_1 and F_2 and verified in the following $F_{2:3}$ generation using SSR markers. We also performed QTL mapping for fiber quality and yield related traits with these three generations. New stable QTLs for fiber quality and yield traits were identified and validated in multiple generations. Our study lays a foundation for fine mapping of fiber quality QTLs and using them in breeding via marker assisted selection (MAS).

Materials and Methods

Materials

CSSL population was constructed by crossing and backcrossing between donor parent Hail and the recurrent parent CCRI36. Hail was a commercially grown cultivar of G. barbadense with a dominant glandless gene, highly resistant to Verticillium wilt, and has excellent fiber quality [42]. While CCRI36 was a widely grown upland cotton cultivar with high yield and early maturity, and developed by the Institute of Cotton Research of Chinese Academy of Agricultural Sciences (State Approval Certificate of Cotton 990007). The development of the CSSL population was reported by Shi et al. [38]. The four BC₅F₄ introgression lines, MBI9804, MBI9855, MBI9752, and MBI9134, with stable and excellent fiber quality performance, were selected as parental lines to construct a double-crossed population of (MBI9804×MBI9855) × (MBI9752×MBI9134). In 2012, the double-crossed F_1 was planted in the experimental farm (Anyang, Henan Province) of the Institute of Cotton Research of Chinese Academy of Agricultural Sciences. All parental lines were planted in two rows and a total of 868 individuals of double-crossed F₁ were planted in 45 rows. Each row was 5m long and 0.8m apart with 20 plants. The F_1 plants were selfed and their fiber and seeds were harvested by individual. In 2013, A total of 839 F₂ individual plants were randomly selected from mixture seeds of all F₁ plants in Anyang in 45 rows, with row length of 5 m, row spacing of 0.8 m, and plant spacing of 0.25 m. In 2014, 237 F_{2:3} families were randomly selected with the different fiber quality and planted in Shihezi, Xinjiang Autonomous Region, in two-narrow-row plots, with row length of 3 m and plant spacing of 0.12 m. The plastic-membrane covering technique and a wide/narrow row spacing pattern were used. Row spacing alternation was 0.2 m and 0.6 m.

Investigation of Fiber Yield and Quality Traits

In 2012 and 2013, the phenotypic traits of each plant were investigated. Naturally opened bolls per plant were harvested for indoor testing, including seed cotton weight, fiber weight, boll weight (BW), lint percentage (LP), fiber length (FL), fiber uniformity (FU), fiber micronaire (FM), and fiber strength (FS). In 2014, 30 naturally opened bolls for each plot were harvested for indoor testing same as those in 2012 and 2013. The fiber quality traits were tested with HFT9000 in the Cotton Quality Supervision and Testing Center of the Ministry of Agriculture of China. HVICC international calibration cotton samples were used.

DNA Extraction and SSR Molecular Detection

In 2012 and 2013, young leaves of the parental lines and the double-crossed F_1 and F_2 individual plants were sampled. DNA was extracted using a modified CTAB method [43]. SSR amplification and polyacrylamide gel electrophoresis were performed following Zhang's description [44]. Based on the previously constructed CCRI36×Hai1 BC₁F₁ genetic linkage map [38], SSR markers were selected at a distance of 10–20 cM. A total of 526 pairs of polymorphic SSR markers between CCRI36 and Hai1 were selected for screening of polymorphisms among the four parental lines. Finally, 51 out of the 526 markers were identified to be polymorphic for genotyping of the double-crossed F_1 and F_2 individual plants. The sequences of the SSR primers were uploaded to the CMD database (<u>http://www.cottonmarker.org/</u>). The primers used in the present study were synthesized by Beijing Sunbiotech Co., Ltd. (Beijing, China).

Analysis of Phenotypic Traits

EXCEL 2013 software was used for the descriptive statistical analysis of fiber quality traits (including FL, FS, FM, and FU) and yield related traits (including BW and LP) for the doublecrossed F_1 and F_2 individual plants and the $F_{2:3}$ family lines. The statistical values included average, maximum, minimum, the transgressive rate over the recurrent parent (%), coefficient of variation, skewness and kurtosis. Correlation analysis and ANOVA were performed using the SPSS20.0 software.

Genotypic Analysis for Parents and Population

Genotypic analysis of the parental lines and population was performed based on the SSR polymorphic results using the GGT2.0 software developed by van Berloo (<u>http://www.</u> <u>plantbreeding.wur.nl/UK/software_ggt.html</u>) [45]. The background recovery rate of the CSSLs to the recurrent parent, number of introgressed segments, and length were calculated.

QTL mapping

The linkage map was constructed using the MapChart2.2 software [46]. QTL mapping was performed using the QTL IciMappingV4.0 software developed by Wang et al. [47]. The nomenclature of QTL was: q + trait abbreviation + chromosome number + serial number of the marker closely linked to the trait. For example, qFS-2-7 represented a QTL controlling fiber strength near the seventh marker on chromosome 2 (Chr2).

Results

Fiber Quality and Yield Traits and Correlation Analysis

The four parents had longer fiber than the recurrent parent CCRI36 (P < 0.05). MBI9804 and MBI9134 had significant stronger fiber than that of CCRI36 (P < 0.01) (Tables 1 and 2). In the three generations, except for FU in 2013 and FM in 2014, the average values of the other traits were higher than those of the recurrent parent CCRI36 (P < 0.05). The absolute value of the skewness was <1, indicating that the fiber quality and yield traits showed a normal distribution in the three generations. The recovery rates to the recurrent parent CCRI36 for FL, FU, FM, and FS increased from F₁ to F_{2:3}, whereas those of BW and LP decreased. The coefficient of variation indicated that fiber quality traits, FM, FS, and FL were highly variable compared to FU.

Correlation analysis between fiber quality and yield traits (<u>Table 3</u>) showed that FS was significantly positively correlated with FL in all three generations, significantly positively correlated with FU in F_1 and F_2 . FL was significantly positively correlated with FU in F_1 and F_2 , significantly negatively correlated with FM in F_2 and $F_{2:3}$, whereas significantly positively correlated were FM in F_1 .

Genotypic Analysis of the Parents and Populations

Introgressed Hai1 chromosome segments in the four parental lines were identified by SSR markers using the GGT2.0 software (<u>S1 Table</u>, <u>Fig 1</u>). The background recovery rate in the four parental lines ranged from 94.1% to 97.6%. Each of the parental line harbored 12–20 introgressed Hai1 chromosome segments, spanning a genetic length of 116.3 cM–283.1 cM, and



Parents	Environment	FL(mm)		FU(%)		FM(unit)		FS(cN/tex)		BW(g/boll)		LP(%)	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
CRI36	2012AY	29.14	1.48	85.18	1.56	4.25	0.45	29.53	1.8	5.05	0.71	36.45	4.73
	2013AY	28.29	0.82	85.2	0.9	4.51	0.11	29.9	1.65	5.4	0.63	37.86	1.77
	2014XJ	28.69	0.6	85.25	1.1	4.27	0.38	29.8	1.46	5.17	0.44	40.22	0.01
MBI9804	2012AY	30.83**	1.63	85.90**	1.55	4.10**	0.47	30.45**	1.15	4.8	0.87	34.45	3
	2013AY	29.93*	1.2	85.99	1.17	4.59	0.12	32.79**	1.58	4.95	0.47	32.87	1.24
MBI9855	2012AY	32.22**	1.61	85.6	1.19	3.6	0.45	32.65	1.28	4.80*	0.78	34.50*	4.08
	2013AY	30.99**	0.81	85.64	0.65	3.71**	0.08	33.12	1.47	5.50*	0.35	30.82*	1.44
MBI9752	2012AY	29.34**	1.03	83.9	1.48	4.01	0.43	31.2	1.32	5.05	0.64	30.1	3.51
	2013 AY	29.32*	1.34	86.15	0.95	4.38	0.12	31.65	1.64	4.8	0.53	35.52	2.83
MBI9134	2012AY	30.22**	2.05	84.75**	1.45	4.02**	0.35	30.40**	1.62	4.02	0.68	34.15	6.18
	2013AY	29.84*	1.01	84.85	0.99	4.17	0.17	32.06**	1.75	4.47	0.47	34.3	4.17

Table 1. Phenotypic analysis of parent fiber quality and yield traits.

Note: AY: Anyang; XJ: Xinjiang

** indicates P < 0.01;

* indicates P < 0.05

doi:10.1371/journal.pone.0159101.t001

accounting for 2.4%–5.9% of the total detected genetic length. The introgressed Hail segments were distributed across 21 chromosomes in all four parental lines, mainly on Chr10, Chr11, Chr20 and Chr23. Chr20 had the most introgressed segment number, whereas no introgressed segments were detected on Chr4, Chr19, Chr22, Chr24 and Chr25. The number of

Table 2. Phenotypic analysis of fiber quality and yield traits in the three populations.

Trait	Year	CCRI36				generation								
		Mean	Min.	Max.	SD		Mean	Min.	Max.	SD	Transgressive rate (%)	CV(%)	Skewness	Kurtosis
FL (mm)	2012	29.14	24.65	30.28	1.48	F ₁	29.39**	23.88	33.63	1.49	86.29	5.06	-0.45	0.11
	2013	28.29	27.41	30.14	0.82	F ₂	30.26**	25.67	34.00	1.23	85.94	4.06	-0.01	0.08
	2014	28.69	27.51	29.77	0.60	F _{2:3}	30.43**	27.28	34.08	1.21	94.09	3.99	0.27	0.19
FU(%)	2012	85.18	81.40	86.20	1.56	F ₁	84.48**	88.00	88.00	1.48	72.58	1.76	-0.49	0.18
	2013	85.20	83.50	86.50	0.90	F ₂	85.82	89.20	89.20	1.28	73.30	1.49	-0.73	1.63
	2014	85.25	85.25	83.50	1.10	F _{2:3}	86.06*	89.20	89.20	0.97	73.84	1.13	-0.21	0.34
FM	2012	4.25	2.61	4.33	0.45	F ₁	3.65	2.06	7.18	0.58	39.29	16.03	0.05	3.90
	2013	4.51	3.88	4.92	0.11	F ₂	4.07**	2.71	5.44	0.46	80.70	11.30	-0.29	0.02
	2014	4.27	3.69	4.91	0.38	F _{2:3}	4.00**	2.93	5.09	0.38	84.81	9.44	0.00	-0.16
FS(cN/tex)	2012	29.53	25.20	41.40	1.80	F ₁	29.58**	24.30	34.40	1.71	72.35	5.78	-0.29	-0.27
	2013	29.90	29.00	34.80	1.65	F ₂	32.65**	26.40	41.50	2.21	78.55	6.77	0.21	0.15
	2014	29.80	27.40	32.70	1.46	F _{2:3}	32.68**	26.90	37.50	1.80	90.72	5.52	-0.09	0.42
BW	2012	5.05	2.85	5.20	0.71	F ₁	4.57*	2.57	7.96	0.73	57.83	15.86	0.24	0.24
	2013	5.40	4.09	6.36	0.63	F ₂	4.90**	2.69	8.42	0.70	46.18	14.36	0.39	1.30
	2014	5.17	4.67	6.65	0.44	F _{2:3}	5.16*	3.48	6.24	0.38	41.07	7.44	-0.34	1.25
LP(%)	2012	36.45	19.75	37.24	4.73	F ₁	31.26**	14.22	50.76	0.03	94.70	8.96	-0.16	3.87
	2013	37.86	34.76	39.31	1.77	F ₂	32.38	22.58	57.15	0.03	36.47	10.45	0.38	3.09
	2014	40.22	36.47	41.18	0.01	F _{2:3}	36.79**	27.72	45.66	0.03	18.07	7.07	-0.22	0.82

Note:

** indicates P < 0.01;

* indicates P < 0.05

doi:10.1371/journal.pone.0159101.t002

traits	Population	FL	FU	FM	FS	LP
FU	F1	0.465**				
	F ₂	0.274**				
	F _{2:3}	0.037				
FM	F ₁	0.106**	0.402**			
	F ₂	-0.203**	0.239**			
	F _{2:3}	-0.370**	0.093			
FS	F ₁	0.566**	0.528**	-0.022		
	F ₂	0.504**	0.290**	-0.008		
	F _{2:3}	0.724**	0.117	-0.312**		
LP	F ₁	0.133**	0.155**	0.101**	0.004	
	F ₂	-0.108**	-0.018	0.375**	-0.257**	
	F _{2:3}	-0.213**	-0.004	0.533**	-0.252**	
BW	F ₁	0.327**	0.460**	0.452**	0.320**	-0.055
	F ₂	0.054	0.155**	0.250**	0.125**	-0.133**
	F _{2:3}	-0.049	0.006	0.123	-0.031	-0.131*

Table 3. Correlation between fiber quality and yield traits in the three populations.

Note:

** indicates P < 0.01 (two-sided);</pre>

* indicates P < 0.05 (two-sided).

doi:10.1371/journal.pone.0159101.t003

homozygous introgressed Hail segments were more than that of the heterozygous introgressed segments in three parental lines, except for MBI9855.

Genotyping analysis (S2 Table) showed that the average recovery rates to CCRI36 in F_1 , F_2 and $F_{2:3}$ generations were 96.1%, 96.3% and 96.0%, respectively. The average length of the homozygous introgressed Hai1 segments ranged from 71.8–110.0 cM in the three generations, whereas the average length of the heterozygous introgressed Hai1 segments was 68.5–103.5 cM. The average length of the total introgressed Hai1 segments was 172.0–182.8 cM.

In all three generations, most individual plants contained 4 or more introgressed Hai1 segments, whereas a few plants contained 1–3 introgressed Hai1 segments. The minimum number of introgressed Hai1 segments was one and the maximum was 27, with an average of 13.6–14.5 introgressed Hai1 segments. The length of the introgressed segments was mainly between 126 cM and 246 cM. In F_2 , no heterozygous segments were detected in three plants. In F_1 , no

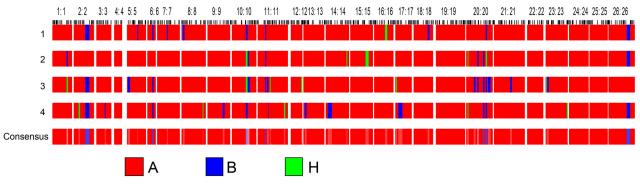


Fig 1. Introgressed Hai1 chromosome segments in the parents. 1: MBI9804; 2: MBI9855; 3: MBI9752; and 4: MBI9134; A: genome of recurrent parent CCRI36; B: homologous Hai1 segments; H: heterozygous Hai1segments.

doi:10.1371/journal.pone.0159101.g001

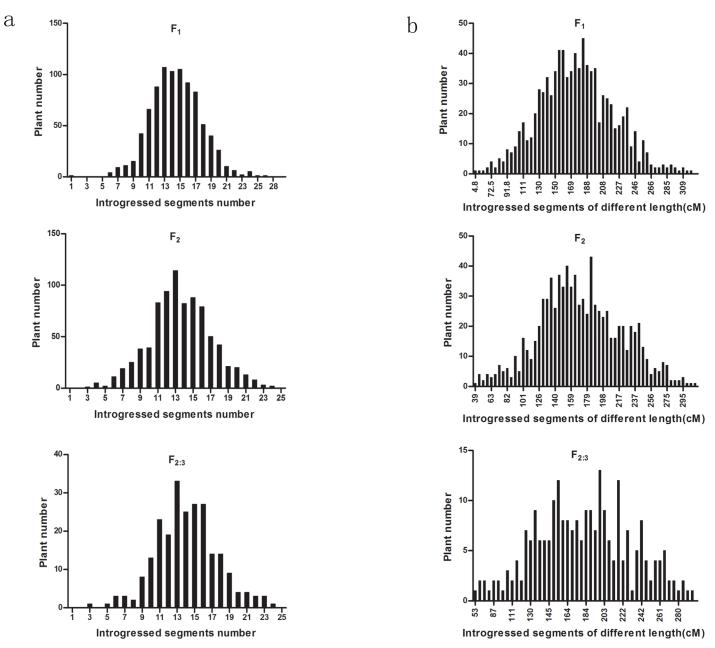


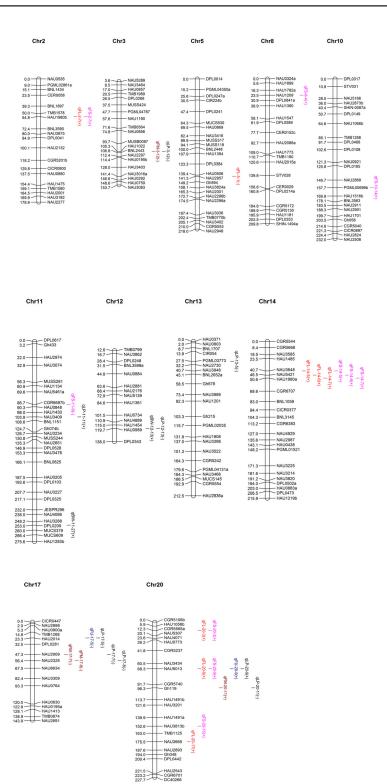
Fig 2. Numbers and total length of the introgressed Hai1 segments in the three generations. a: Number of the introgressed Hai1 segments; b: Total length of the introgressed Hai1 segments.

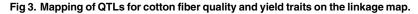
doi:10.1371/journal.pone.0159101.g002

homozygous segments were detected in four plants and one introgressed segment was detected in only one plant (<u>S3 Table</u>, Fig 2a and 2b).

QTL Mapping

A total of 35 QTLs for the yield and fiber quality traits was identified in F_1 , F_2 , and $F_{2:3}$, including 24 QTLs controlling four cotton fiber quality traits and 11 QTLs controlling two cotton yield traits (<u>S4 Table</u>, Fig 3). These QTLs were distributed across 11 chromosomes, with LOD values between 2.31 and 26.38, collectively explained 1.78%–20.27% of the observed





doi:10.1371/journal.pone.0159101.g003

phenotypic variations. Three QTLs were detected in 3 generations, and the 13 QTLs were detected in two generations, these 16 QTLs were regarded as stable ones, of which included 12 fiber quality-related QTLs and 4 yield-related QTLs.

Fiber length: Eight QTLs controlling FL were detected on four chromosomes (Chr2, Chr14, Chr5 and Chr20), explaining 1.93%–19.68% of the observed phenotypic variations. Among these QTLs, one on Chr2 (qFL-2–7) and three on Chr14 (qFL-14-5, qFL-14-6 and qFL-14-7) were detected in both F_1 and F_2 generations. Three QTLs in $F_{2:3}$ were mapped on Chr5 (qFL-5-1) and Chr20 (qFL-20-3 and qFL-20-17), explaining 4.91%, 5.37%, and 19.68% of the observed phenotypic variations, respectively. The negative additive effect for qFL-5-1 indicated that CCRI36 alleles increased fiber length.

Fiber strength: A total of 10 QTLs for FS were detected on six chromosomes (Chr2, Chr8, Chr10, Chr11, Chr14 and Chr20), explaining 1.84%-6.49% of the observed phenotypic variations. Among these QTLs, four that were detected in both F₁ and F₂ were mapped on Chr2 (qFS-2-7) and Chr14 (qFS-14-5, qFS-14-6 and qFS-14-7), explaining 2.07%-5.77%, 5.53%-6.49%, 5.06%-6.06%, and 5.24%-5.78% of the observed phenotypic variations, respectively. One QTL mapped on Chr20 was detected in F_{2:3} (qFS-20-3), explaining 4.73% of the observed phenotypic variations. Of the remaining QTLs, one was detected in F₁ and four in F₂.

Micronaire: Four QTLs for FM were mapped on three chromosomes (Chr3, Chr17 and Chr20), explaining 1.95%–11.69% of the observed phenotypic variations. Among these QTLs, three were detected both in F_2 and $F_{2:3}$, and mapped on Chr3 (qFM-3-12) and Chr17 (qFM-17-7 and qFM-17-8), explaining 8.86%–10.46%, 5.12%–11.69% and 6.78%–7.67% of the observed phenotypic variations, respectively. The negative additive effects indicated that CCRI36 alleles increased micronaire value.

Fiber uniformity: Two QTLs for fiber uniformity were mapped on Chr17 and Chr20. qFU-17-5 explained 2.49%–3.83% of the observed phenotypic variations with a negative additive effect. QTL qFU-20-9 was detected in both F_1 and F_2 , explaining 3.57% of the observed phenotypic variations, with a positive additive effect which indicated Hai1 alleles increased fiber uniformity.

Boll weight: Two QTLs for BW were mapped on Chr11 and Chr12, explaining 6.02%– 9.50% of the observed phenotypic variations. The positive additive effects indicated that Hai1 alleles increased boll weight.

Lint percentage: Nine QTLs for LP were mapped on 6 chromosomes (Chr3, Chr11, Chr12, Chr13, Chr17 and Chr20), explaining 1.84%–13.50% of the observed phenotypic variations. Three QTLs, qLP-3-12, qLP-17-7 and qLP-17-8, were detected in all F_1 , F_2 and $F_{2:3}$, explaining 8.01%–9.91%, 5.28%–13.5% and 4.4%–11.57% of the observed phenotypic variations, respectively. qLP-12-13 was detected in both F_2 and $F_{2:3}$, and explained 5.05%–10.77% of the observed phenotypic variations. The negative additive effects indicated that CCRI36 alleles increased lint percentage.

Discussion

Assessment and Application of Chromosome Segment Substitution Lines

CSSLs are usually applied to investigate the genetic behavior and effects of chromosome introgression **Segment** from the donor parent in the background of the recurrent parent. Its application for QTL mapping generally improves accuracy. CSSLs also provide permanent segregation populations for studying multi-environmental stability of the mapped QTLs. In the previous studies, a set of CSSL population was constructed using a widely planted upland cotton cultivar, CCRI36 as the recurrent parent, which is characterized of high yield and early maturity, and a sea island cotton Hai1 as the donor parent, which had good fiber quality and a high level of resistance to *Verticillium* wilt [38, 39, 40, 41]. In the current study, four introgression lines with excellent fiber quality which derived from this CSSL population were used to construct the segregating populations of double-crossed F_1 , F_2 , and $F_{2:3}$. The introgressed Hai1 Chromosome segments were detected in individual plants of all three generations. The genotyping of each individual plant in the three generations was relatively clear. The recovery rate to the recurrent parent in all three generations was >95%. These introgression lines are ideal materials for further fine mapping, gene interaction analysis, heterosis mechanism study, and functional genomics.

Comparison of Our Results with the Previous Findings

The QTLs detected in the present study, including 35 QTLs controlling fiber quality traits and 11 QTLs controlling yield traits, were mapped to 11 chromosomes (Chr2, Chr3, Chr5, Chr8, Chr10, Chr11, Chr12, Chr13, Chr14, Chr17, and Chr20). Among these QTLs, three (qLP-17-7, qLP-17-8, and qLP-3-12) were consistently detected in all three generations, and 13 were consistently detected in two generations (qFL-14-5, qFL-14-6, qFL-14-7, qFL-2-7, qFS-14-5, qFS-14-6, qFS-14-7, qFS-2-7, and qFU-20-9 detected in F1 and F2; qFM-17-7, qFM-17-8, qFM-3-12, and qLP-12-13 detected in F_2 and $F_{2:3}$ (S4 Table). Fifteen QTLs detected in the present study were reported in previous studies, including qFL-2-7, qFS-2-7, qBW-11-27, qBW-12-13, qLP-12-13, qFL-14-5, qFS-14-5, qFS-14-5, qFM-17-7, qFM-17-7, qFU-17-5, qLP-17-5, qFL-20-3, qFS-20-3, and qFS-20-15 $\left[\frac{40-41}{48}, \frac{48-50}{50}\right]$, based on the common markers harbored in the same confidence intervals of the same chromosome positions where these QTLs were mapped. Eight QTLs (qLP-17-8, qLP-3-12, qFL-14-6, qFL-14-7, qFS-14-6, qFS-14-7, qFM-17-8 and qFM-3-12) are reckoned to be novel stable QTLs. The detection of QTLs in multiple generations and different genetic backgrounds was suggestive of the stabilities of the genetic effects. The special attentions should be paid to these stable QTLs because the quantitative traits are usually susceptible to the various environmental factors, and the stable QTLs might increase reliability and efficiency of selection and play important roles in fiber yield and quality improvement via MAS [51-56].

QTL Cluster and Linkage Distribution

Cluster distribution of QTLs is a relatively common phenomenon [14, 17, 53, 55–58]. Said et al. [59] comprehensively analyzed 2,134 previously reported QTLs in intra- and inter-species populations and detected numerous QTLs, which were distributed in clusters in certain chromosome regions in the specific populations. In the present study, QTLs controlling different traits were also detected at the same SSR marker locus. QTLs for FL and FS were mapped to the neighboring regions of markers HAU1980b on Chr2, NAU3648, NAU5421 and HAU1980a on Chr14, and CGR5565a, NAU5013, and NAU3665 on Chr20. QTLs for FU and LP were mapped to the neighboring regions of marker NAU5013 (S4 Table). QTLs for FM (qFM3-12, qFM-17-7, and qFM-17-8) were consistently detected in the neighboring regions of the molecular markers linked to the QTLs for LP (qLP-3-12, qLP-17-7, and qLP-17-8), suggesting that these major QTLs are closely linked to the same markers in the introgressed Hai1 segment, thereby providing an explanation for the significant correlations between the two traits in all three generations (Table 3). These results indicate that these loci might function as pleiotropic genes or are closely linked to the various other genes.

The chromosome segments of these QTL hotspot clusters could be useful for molecular breeding based on common molecular markers [59]. When the chromosome segments clustered both the favorable alleles of the QTLs for cotton fiber quality and yield traits, it could be

more easily used for simultaneous improvement of traits. However, when the chromosome segments clustered the negatively correlated favorable alleles of QTLs, it would be very difficult to simultaneously improve these traits. An in-depth study of this linkage mechanism and breaking the linkage between cumbersome genes would play a significant role in cotton molecular breeding.

Sources of QTL Synergistic Genes

Among the QTLs for fiber quality traits mapped in the present study, the synergistic genes for 14 QTLs were from CCRI36, whereas the synergistic genes for 21 QTLs were from Hai1. Among the QTLs controlling cotton yield traits, the synergistic genes for 8 QTLs were from CCRI36, whereas the synergistic genes for 3 QTLs were from Hai1. These results suggest that fiber quality or yield QTLs are not necessarily derived from the superior parent and the parent with relatively poor traits can also contribute genes that favor fiber yield and quality. Our findings also indicate that introgression between upland cotton and sea island cotton may broaden genetic variations as well as increase the potential of favorable gene rearrangements.

Supporting Information

S1 Table. Statistical analysis of Hai1 introgression segments in the parental lines. (XLSX)

S2 Table. Statistical analysis of Hai1 introgressed segment length in the three generations. (XLSX)

S3 Table. Statistical analysis of Hai1 substitution segments in the three generations. (XLSX)

S4 Table. QTLs related to fiber quality and yield traits detected in the three generations. (XLSX)

Author Contributions

Conceived and designed the experiments: YLY YZS.

Performed the experiments: HCZ YNT.

Analyzed the data: HCZ WKG YLY YZS.

Contributed reagents/materials/analysis tools: WKG AYL JWL JWG QG WWS ZYD LLK HHS TTC.

Wrote the paper: HCZ WKG.

References

- 1. Smith CW, Coyle GG. Association of Fiber Quality Parameters and Within-Boll Yield Components in Upland Cotton. Crop Science. 1997; 37(6): 1775–1779.
- Liu JY, Zhao GR, Li J. Molecular engineering on quality improvement of cotton fiber. Acta Botanica Sinica. 2000; 42(10): 991–995.
- Clement JD, Constable GA, Stiller WN, Liu SM. Negative associations still exist between yield and fiber quality in cotton breeding programs in Australia and USA. Field Crops Research. 2012; 128: 1–7.
- 4. Percy RG, Cantrell RG, Zhang J. Genetic variation for agronomic and fiber properties in an introgressed recombinant inbred population of cotton. Crop Science. 2006; 46(3): 1311–1317.

- Reinisch AJ, Dong JM, Brubaker CL, Stelly DM, Wendel JF, Paterson AH. A detailed RFLP map of cotton, Gossypium hirsutum × Gossypium barbadense: chromosome organization and evolution in a disomic polyploid genome. Genetics. 1994; 138(3): 829–847. PMID: 7851778
- 6. Wang K, Wang Z, Li F, Ye W, Wang J, Song G, et al. The draft genome of a diploid cotton *Gossypium* raimondii. Nature Genetics; 2012; 44(10): 1098–1103. doi: <u>10.1038/ng.2371</u> PMID: <u>22922876</u>
- 7. Li FG, Fan GY, Wang KB, Sun FM, Yuan YL, Song GL, et al. Genome sequence of the cultivated cotton *Gossypium arboretum*; 2014; 46(6): 567–572.
- Zhang TZ, Hu Y, Jiang WK, Fang L, Guan XY, Chen JD, et al. Sequencing of allotetraploid cotton (Gossypium hirsutum L. acc. TM-1) provides a resource for fiber improvement. Nature Biotechnology; 2015; 33(5): 531–U252. doi: 10.1038/nbt.3207 PMID: 25893781
- Li FG, Fan GY, Lu CR, Xiao GH, Zou CS, Kohel RJ, et al. Genome sequence of cultivated Upland cotton (*Gossypium hirsutum* TM-1) provides insights into genome evolution. Nature Biotechnology; 2015; 33(5): 524–U242. doi: <u>10.1038/nbt.3208</u> PMID: <u>25893780</u>
- Paterson AH, Saranga Y, Menz M, Jiang CX, Wright RJ. QTL analysis of genotype × environment interactions affecting cotton fiber quality. Theoretical and Applied Genetics. 2003; 106(3): 384–396. PMID: 12589538
- Jiang CX, Wright RJ, El-Zik KM, Paterson AH. Polyploid formation created unique avenues for response to selection in *Gossypium* (cotton). Proceedings of the National Academy of Sciences of the United States of America. 1998; 95(8): 4419–4424. PMID: <u>9539752</u>
- 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*). Theoretical and Applied Genetics. 2004; 108(2): 280–291. PMID: <u>14513220</u>
- Song XL, Wang K, Guo WZ, Zhang J, Zhang TZ. A comparison of genetic maps constructed from haploid and BC1 mapping populations from the same crossing between Gossypium hirsutum L. and Gossypium barbadense L. Genome. 2005; 48(3): 378–390. PMID: <u>16121235</u>
- Yu J, Zhang K, Li S, Yu S, Zhai H, Wu M, et al. Mapping quantitative trait loci for lint yield and fiber quality across environments in a *Gossypium hirsutum* × *Gossypium barbadense* backcross inbred line population. Theoretical and Applied Genetics. 2013; 126(1): 275–287. doi: <u>10.1007/s00122-012-1980-x</u> PMID: <u>23064252</u>
- 15. Shen X, Guo W, Lu Q, Zhu X, Yuan Y, Zhang T. Genetic mapping of quantitative trait loci for fiber quality and yield trait by RIL approach in Upland cotton. Euphytica. 2007; 155(3): 371–380.
- Lacape JM, Llewellyn D, Jacobs J, Arioli T, Becker D, Calhoun S, et al. Meta-analysis of cotton fiber quality QTLs across diverse environments in a *Gossypium hirsutum* × *G. barbadense* RIL population. BMC Plant Biology. 2010; 10(1): 107–113.
- 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. × *G barbadense* L.) RIL population. Molecular Genetics and Genomics. 2014; 289(6): 1347–1367. doi: 10.1007/s00438-014-0930-5 PMID: 25314923
- Zhang Z, Li J, Muhammad J, Cai J, Jia F, Shi Y, et al. High resolution consensus mapping of quantitative trait loci for fiber strength, length and micronaire on chromosome 25 of the upland cotton (*Gossypium hirsutum* L.). Plos One, 2015; 10(8): e0135430. doi: <u>10.1371/journal.pone.0135430</u> PMID: <u>26262992</u>
- 19. Eshed Y, Zamir D. A genomic library of *Lycopersicon pennellii* in *L. esculentum*: a tool for fine mapping of genes. Euphytica. 1994; 79(3): 175–179.
- Eshed Y, Zamir D. An introgression line population of *Lycopersicon pennellii* in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. Genetics. 1995; 141(3): 1147–1162. PMID: 8582620
- Wan X, Weng J, Zhai H, Wang J, Lei C, Liu X, et al. Quantitative trait loci (QTL) analysis for rice grain width and fine mapping of an identified QTL allele gw-5 in a recombination hotspot region on chromosome 5. Genetics. 2008; 179(4): 2239–2252. doi: 10.1534/genetics.108.089862 PMID: 18689882
- Liu Q, Qin J, Li T, Liu E, Fan D, Edzesi WM, et al. Fine Mapping and Candidate Gene Analysis of qSTL3, a Stigma Length-Conditioning Locus in Rice (Oryza sativa L.). PLoS One. 2015; 10(6): e0127938. doi: <u>10.1371/journal.pone.0127938</u> PMID: <u>26030903</u>
- Li X, Wang W, Wang Z, Li K, Lim YP, Piao Z. Construction of chromosome segment substitution lines enables QTL mapping for flowering and morphological traits in *Brassica rapa*. Plant Science. 2015; 6: 432.
- Stelly DM, Saha S, Raska DA, Jenkins JN, McCarty JC, Gutierrez OA. Registration of 17 upland (Gossypium hirsutum) cotton germplasm lines disomic for different G. barbadense chromosome or arm substitutions. Crop Science. 2005; 45(6): 2663–2665.

- Saha S, Jenkins JN, Wu JX, McCarty JC, Gutierrez OA, Percy RG, et al. Effects of chromosome-specific introgression in upland cotton on fiber and agronomic traits. Genetics. 2006; 172(3): 1927–1938. PMID: 16387867
- Saha S, Jenkins JN, Wu J, McCarty JC, Stelly DM. Genetic analysis of agronomic and fiber traits using four interspecific chromosome substitution lines in cotton. Plant Breeding. 2008; 127(6): 612–618.
- Saha S, Wu J, Jenkins JN, McCarty JC, Hayes R, Stelly DM. Genetic dissection of chromosome substitution lines of cotton to discover novel *Gossypium barbadense* L. alleles for improvement of agronomic traits. Theoretical and Applied Genetics. 2010; 120(6):1193–205. doi: <u>10.1007/s00122-009-1247-3</u> PMID: 20094702
- McCarty JC, Wu J, Saha S, Jenkins JN, Hayes R. Effects of chromosome 5sh from Gossypium barbadense L. on flower production in G. hirsutum L. Euphytica. 2006; 152(1):99–107.
- Jenkins JN, McCarty JC, Wu J, Saha S, Gutierrez O, Hayes R, et al. Genetic effects of thirteen Gossypium barbadense L. chromosome substitution lines in topcrosses with upland cotton cultivars: I. Yield and yield components. Crop Science. 2006; 46(3): 1169–1178.
- Jenkins JN, McCarty JC, Wu J, Saha S, Gutierrez O, Hayes R, et al. Genetic effects of thirteen Gossypium barbadense L. chromosome substitution lines in topcrosses with upland cotton cultivars: II. Fiber quality traits. Crop Science. 2007; 47(2): 561–572.
- Jenkins JN, McCarty JC, Wu J, Hayes R, Stelly D (2011). Genetic effects of nine Gossypium barbadense L. chromosome substitution lines in top crosses with five elite Upland cotton G. hirsutum L. cultivars. Euphytica. 2012; 187(2): 161–173.
- Zhang J, Percy R, McCarty J Jr. Introgression genetics and breeding between Upland and Pima cotton: a review. Euphytica. 2014; 198(1):1–12.
- Wei W. Genetic Evaluation of Introgression Line 6 and Identification of the Introgressed Segments. Cotton Science. 2009; 21(5): 394–398.
- Zhu Y, Wang P, Guo W, Zhang T. Mapping QTLs for lint percentage and seed index using Gossypium barbadense chromosome segment introgression lines. Acta Agronomica Sinica. 2010; 36(8): 1318–1323.
- Wang P, Zhu Y, Song X, Cao Z, Ding Y, Liu B, et al. Inheritance of long staple fiber quality traits of Gossypium barbadense in G. hirsutum background using CSILs. Theoretical and Applied Genetics. 2012; 124(8): 1415–1428. doi: 10.1007/s00122-012-1797-7 PMID: 22297564
- Cao Z, Wang P, Zhu X, Chen H, Zhang T. Ssr marker-assisted improvement of fiber qualities in Gossypium hirsutum using G. barbadense introgression lines. Theoretical and Applied Genetics. 2014; 127(3): 587–594. doi: 10.1007/s00122-013-2241-3 PMID: 24306319
- 37. Wang B, Nie Y, Lin Z, Zhang X, Liu J, Bai J. Molecular diversity, genomic constitution, and QTL mapping of fiber quality by mapped SSRs in introgression lines derived from *Gossypium hirsutum* × *G. darwinii watt*. Theoretical and Applied Genetics. 2012; 125(6): 1263–1274. doi: <u>10.1007/s00122-012-1911-x PMID</u>: 22733444
- Shi Y, Li W, Li A, Ge R, Zhang B, Li J, et al. Constructing a high-density linkage map for Gossypium hirsutum × Gossypium barbadense and identifying QTLs for lint percentage. Journal of Integrative Plant Biology. 2015; 57(5): 450–467. doi: 10.1111/jipb.12288 PMID: 25263268
- Liang Y, Jia Y, Li A, Zhang B, Liu G, Li J, et al. Phenotyping traits related to yield and quality of BC₅F₂ substitution lines in cotton and their QTL mapping. Molecular Plant Breeding. 2010; 8(2): 221–230.
- Zhang J, Shi Y, Liang Y, Jia Y, Zhang B, Li J, et al. Evaluation of Yield and Fiber Quality Traits of Chromosome Segment Substitution Lines Population (BC₅F₃ and BC₅F_{3:4}) in Cotton. Journal of Plant Genetic Resources. 2012; 13(5): 773–781.
- **41.** He R, Shi Y, Zhang J, Liang Y, Zhang B, Li J, et al. QTL mapping for plant height using chromosome segment substitution lines in upland cotton. Acta Agronomica Sinica. 2014; 40(3):457–465.
- Sun Q, Cai YF, Xie YF, Mo JC, Youlu Y, Shi YZ, et al. Gene expression profiling during gland morphogenesis of a mutant and a glandless upland cotton. Molecular Biology Reports. 2010; 37(7): 3319–3325. doi: <u>10.1007/s11033-009-9918-3</u> PMID: <u>19888674</u>
- Paterson AH, Brubaker CL, Wendel JF. A rapid method for extraction of cotton (Gossypium spp.) genomic DNA suitable for RFLP or PCR analysis. Plant Molecular Biology Reporter. 1993; 11(2): 122–127.
- Zhang J, Wu YT, Guo WZ, Zhang TZ. Fast Screening of Microsatellite Markers in Cotton with PAGE/silver Staining. Acta Gossypii Sinica. 2000; 12(5): 267–269.
- van Berloo R. GGT 2.0: Versatile software for visualization and analysis of genetic data. Journal of Heredity. 2008; 99(2): 232–236. doi: <u>10.1093/jhered/esm109</u> PMID: <u>18222930</u>
- Voorrips RE. Map Chart: Software for the graphical presentation of linkage maps and QTLs. Journal of Heredity. 2002; 93(1): 77–78. PMID: <u>12011185</u>

- 47. Wang J, Wan X, Crossa J, Crouch J, Weng J, Zhai H, et al. QTL mapping of grain length in rice (*Oryza sativa* L.) using chromosome segment substitution lines. Genetical Research. 2006; 88(2): 93–104. PMID: <u>17125584</u>
- Ma L, Shi Y, Lan M, Yang Z, Zhang J, Zhang B, et al. Evaluation on fiber yield and quality traits of chromosome segment substitution lines from *Gossypium hirsutum* × *Gossypium barbadense*. Cotton Science. 2013; 25(6): 486–495.
- Xu Z, Zhang C, Ge X, Wang N, Zhou K, Yang X, et al. Construction of a high-density linkage map and mapping quantitative trait loci for somatic embryogenesis using leaf petioles as explants in upland cotton (*Gossypium hirsutum* L.). Plant Cell Reports. 2015; 34(7): 1177–1187. doi: <u>10.1007/s00299-015-1776-y</u> PMID: <u>25758337</u>
- **50.** Li L. Molecular marker study on Cotton Fiber Quality, yield and resistance to verticillium wilt using Advanced Backcross recombinational lines between Upland Cotton (*G. hirsutum*. L) and Island Cotton (*G. barbadense*. L). Chinese Academy of Agriculture Sciences Master Dissertation. 2008.
- Lacape JM, Nguyen TB. Mapping quantitative trait loci associated with leaf and stem pubescence in cotton. Journal of Heredity. 2005; 96(4):441–4. PMID: <u>15829730</u>
- Shen X, Guo W, Lu Q, Zhu X, Yuan Y, Zhang T. Genetic mapping of quantitative trait loci for fiber quality and yield trait by RIL approach in Upland cotton. Euphytica. 2007; 155(3):371–80.
- 53. Zhang Z-S, Hu M-C, Zhang J, Liu D-J, Zheng J, Zhang K, et al. Construction of a comprehensive PCRbased marker linkage map and QTL mapping for fiber quality traits in upland cotton (*Gossypium hirsutum* L.). Molecular Breeding. 2009; 24(1):49–61.
- Zhang K, Zhang J, Ma J, Tang S, Liu D, Teng Z, 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.). Molecular Breeding. 2012; 29(2):335–48.
- Tang S, Teng Z, Zhai T, Fang X, Liu F, Liu D, 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.
- 56. Wang H, Huang C, Guo H, Li X, Zhao W, Dai B, et al. QTL Mapping for Fiber and Yield Traits in Upland Cotton under Multiple Environments. PLoS ONE. 2015; 10(6): e0130742. PMID: <u>26110526</u>
- 57. 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: 143–156.
- Said JI, Lin Z, Zhang X, Song M, Zhang J. 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: 776. doi: 10.1186/1471-2164-14-776 PMID: 24215677
- 59. Said J, Song M, Wang H, Lin Z, Zhang X, Fang D, et al. A comparative meta-analysis of QTL between intraspecific Gossypium hirsutum and interspecific G. hirsutum × G. barbadense populations. Molecular Genetics and Genomics. 2015; 290(3):1003–25. doi: 10.1007/s00438-014-0963-9 PMID: 25501533