



# Association analysis of yield and fiber quality traits in *Gossypium barbadense* with SSRs and SRAPs

X.Q. Wang<sup>1</sup>, Y. Yu<sup>1,2</sup>, W. Li<sup>1</sup>, H.L. Guo<sup>1</sup>, Z.X. Lin<sup>1</sup> and X.L. Zhang<sup>1</sup>

<sup>1</sup>National Key Laboratory of Crop Genetic Improvement & National Centre of Plant Gene Research, Huazhong Agricultural University, Wuhan, Hubei, China

<sup>2</sup>Cotton Institute, Xinjiang Academy of Agriculture and Reclamation Science, Shihezi, Xinjiang, China

Corresponding author: Z.X. Lin  
E-mail: linzhongxu@mail.hzau.edu.cn

Genet. Mol. Res. 12 (3): 3353-3362 (2013)

Received September 11, 2012

Accepted January 2, 2013

Published September 4, 2013

DOI <http://dx.doi.org/10.4238/2013.September.4.1>

**ABSTRACT.** Cotton is an important cash crop. Mining for quantitative trait loci related to yield and fiber quality traits using association analysis has many advantages for cotton research. In this study, 170 simple sequence repeats (SSRs) and 258 sequence-related amplified polymorphisms (SRAPs) were used to analyze the association of 3 yield component traits and 5 fiber quality traits of 55 *Gossypium barbadense* accessions in 2009 and 2010. Principal component analysis of SSRs and SRAPs showed 3 and 2 subgroups, respectively. The boundaries between the SRAP groups were much more defined than those of the SSRs. A mixed linear model was used to analyze association of yield and fiber quality traits with SSRs and SRAPs. A total of 72 loci were detected, including 28 loci of SSRs and 44 loci of SRAPs; 26 of these loci were related to yield component traits, and 46 of these loci were related to fiber quality traits. The mean phenotypic variations explained in the SSR and SRAP analysis were 8.89 and 8.61%, respectively. The locus with the highest phenotypic variation explained was NAU1164

(23.33%), which was related to fiber uniformity. The comparison of association results between the two datasets showed that mining quantitative trait loci using association analysis was more efficient with SRAPs than with SSRs.

**Key words:** *Gossypium barbadense*; Association analysis; SSR; SRAP

## INTRODUCTION

Cotton is one of the most important natural fiber crops in the world. Among the 4 cultivated species (*Gossypium hirsutum*, *G. barbadense*, *G. arboreum*, *G. herbaceum*), *G. barbadense* has the most ideal fiber quality traits, despite its lower yield relative to *G. hirsutum*. Therefore, more attention should be paid to *G. barbadense*, which offers great potential for yield production and the development of the textile industry.

Most traits in plants, such as yield components and quality, are quantitative traits. Currently, there are 2 strategies to search for quantitative trait loci (QTLs): the traditional method based on the linkage map, and a newer method based on linkage disequilibrium (LD). As linkage map construction requires the selection of appropriate parents, and then the growth of temporary or permanent populations, it is a particularly time-consuming process. More importantly, the polymorphisms in such populations will be relatively low as the alleles are limited to 1 locus of the 2 parents (Salvi and Tuberosa, 2005).

Recently, with the extensive development of simple sequence repeats (SSRs) and single nucleotide polymorphisms, association analysis has been widely adopted for mining QTLs in genome research. As a robust tool to detect QTLs, the advantages of association mapping are obvious. Natural populations can be used in association mapping, with the exception of some specially constructed populations such as nested association mapping or multi-parent advanced generation inter-cross populations. More allele variants can be detected from the widely collected germplasm, and the full recombination of the population promotes higher accuracy in association mapping than is possible in linkage mapping (Flint-Garcia et al., 2003; Yang et al., 2007). In association analysis based on LD, many potential correlative factors can cause errors, including: the mating system (LD attenuation of inbred species is slower than in outcrossed species), bottleneck events that can increase LD, and migration among populations, which can also increase LD. In particular, gene flow among subgroups can substantially enhance LD in the whole population, resulting in false associations (Rafalski and Morgante, 2004; Stich et al., 2005).

In order to reduce the probability of false association, control of population structure is necessary in association analysis. Some studies have used principal component analysis (PCA) results to divide groups (Yan et al., 2009; Wen et al., 2012), while others rely on the mean log probability of the data [LnP(D)] obtained from the STRUCTURE software (Abdurakhmonov et al., 2008; Yao et al., 2009). The most commonly used models for analyzing the association are one-way analysis of variance (ANOVA), multiple linear regressions, general linear models (GLM), and mixed linear models (MLM) (Grapes et al., 2006; Wu et al., 2007; Kantartzi and Stewart, 2008; Abdurakhmonov et al., 2008, 2009). MLM has been widely used more recently, because these models allow for the consideration of population structure and kinship, which reduces the probability of false association.

Since its introduction for quantitative trait analysis in plants, association analysis has been widely used in rice, *Arabidopsis*, maize, and many other species. However, association analysis in cotton has been much rarer due to the complexity of the genome. To date, association analysis of cotton has mostly focused on fiber quality traits, and several models have been applied, including multiple regression models (Wu et al., 2007), MLM (Abdurakhmonov et al., 2008, 2009), and GLM (Kantartzi and Stewart, 2008).

In order to explore QTLs related to yield and to fiber quality traits in *G. barbadense*, SSR and sequence-related amplified polymorphism (SRAP) markers were used to determine the genetic structure of 55 *G. barbadense* accessions collected from various regions throughout the world. Data of yield and fiber quality traits collected in 2009 and 2010 were included in an association analysis using mixed linear models. Furthermore, association results obtained from SSRs and from SRAPs were compared.

## MATERIAL AND METHODS

### Plant material

The materials used in this study were the same as in Li et al. (2008), with the exception of Xinhai5 (Accession No. 37), which was excluded, because it was adverse in determining correct population structure. In consequence, a total of 55 accessions were used in this study.

### Field experiment and data collection

A field experiment was carried out in the Xinjiang region, which has the most suitable climate for *G. barbadense* growth in China. In 2009 and 2010, 2 replicates of 55 accessions were grown in a randomized plot design with a single row plot and 100 individuals per row. Using a specific wide-narrow distance planting pattern, the row spacing was 66 ( $\pm 10$ ) cm, with 9.5 cm between individuals.

Yield and the component traits were measured in mid-September of 2009 and 2010. A total of 80 bolls were collected for determining lint weight per boll (LW), lint percentage (LP), and seed cotton weight per boll (SW).

To measure fiber quality traits, 10-15 g fibers was collected from middle fruit branches and sent to the Cotton Quality Supervision, Inspection and Testing Center, Ministry of Agriculture, Anyang, China. The tests were conducted at a temperature of 20°C, and 65% relative humidity. Variables tested included fiber elongation (FE), and fiber strength (FS), fiber uniformity (FU), fiber upper half mean length (FUHML), micronaire value (MV).

The SPSS 17.0 software (<http://www-01.ibm.com/software/analytics/spss/>) was used to calculate basic statistics (e.g., mean, maximum, minimum, standard value), and statistics to determine whether or not the variables were normally distributed and therefore appropriate for use in linear models (e.g., skew, kurtosis).

### Genotype data

SSR data was obtained from Wang et al. (2011b), and SRAP data was obtained from Li et al. (2008).

## Genetic structure

Based on the raw data, a genetic distance matrix was first constructed using the GenAIEx6.2 software (Peakall and Smouse, 2006). A PCA plot and matrix were then obtained based on the GD matrix.

## Association analysis

Association analysis of yield and fiber quality traits was performed using the TASSEL software (Version 2.1; <http://www.maizegenetics.net>). As the requirements of TASSEL, data from each individual locus were converted into its genotypic form. Genotypic data, filtered for a minimum allele frequency of 5%, were first used to construct a “point” matrix and a kinship matrix, and then the two matrixes were combined with the PCA matrix and phenotypic data to run an MLM.

## RESULTS

### Phenotypic performance

The statistics for the phenotypic data were listed in Table 1. Coefficients of variation of SW (2010), LW (2010) and MV (2009) were all over 10%. The statistical test for skew and kurtosis revealed that all distributions were normally distributed.

**Table 1.** Statistical characteristic of yield and fiber quality traits of *Gossypium barbadense*.

Traits	Year	Max	Min	Mean	Std	CV (%)	Skewness	Test of skewness	Kurtosis	Test of kurtosis
LP	2009	39.90	30.79	35.21	1.78	5.05	0.24	0.72	0.67	0.99
	2010	41.23	29.29	35.32	2.26	6.40	-0.14	0.43	0.70	1.08
LW	2009	1.42	1.01	1.23	0.10	7.90	-0.21	0.64	-0.26	0.39
	2010	1.88	0.88	1.43	0.19	12.96	-0.47	1.44	0.82	1.25
SW	2009	4.13	3.04	3.51	0.27	7.66	0.50	1.49	-0.29	0.44
	2010	5.42	2.65	4.07	0.49	12.14	-0.04	0.11	1.02	1.56
FE	2009	6.50	5.90	6.28	0.16	2.49	-0.52	1.53	-0.28	0.42
	2010	5.60	4.40	5.08	0.29	5.71	-0.17	0.53	-0.73	1.11
FS	2009	44.60	32.40	37.92	2.82	7.45	0.59	1.75	-0.28	0.42
	2010	42.90	30.20	36.04	3.27	9.07	0.73	2.22	-0.45	0.68
FU	2009	89.00	82.90	85.94	1.32	1.54	-0.03	0.08	-0.04	0.06
	2010	88.4	80.20	85.26	1.63	1.91	-0.69	2.10	1.14	1.74
FUHML	2009	38.62	30.56	35.35	1.92	5.43	-0.29	0.86	-0.40	0.60
	2010	37.83	30.11	34.56	2.00	5.80	-0.39	1.20	-0.68	1.04
MV	2009	5.24	3.04	4.23	0.45	10.55	-0.17	0.49	0.78	1.16
	2010	5.25	3.42	4.36	0.41	9.31	-0.08	0.24	-0.28	0.42

CV = coefficient of variation; LP = lint percentage; LW = lint weight per boll; SW = seed weight per boll; FE = fiber elongation; FS = fiber strength; FU = fiber uniformity; FUHML = fiber upper half mean length; MV = micronaire value.

### Correlation analysis of yield and fiber quality traits

Correlations among 3 yield traits and 5 fiber quality traits are listed in Table 2. The following variables were positively correlated (at  $P < 0.01$  significance levels): LW with LP and SW; FU with FUHML and FS; and FUHML with SW and FS. The following variables were negatively correlated (at  $P < 0.01$  significant levels): FS with LP and FE, FU with FE, FUHML with LP, FE, and MV.

**Table 2.** Correlation coefficients among yield and fiber quality traits in 2009 and 2010.

	LP	LW	SW	FE	FS	FU	FUHML	MV
LP	-	0.40**	-0.08	0.31*	-0.48**	-0.33*	-0.50**	0.25
LW	0.38**	-	0.88**	-0.45**	0.29*	0.20	0.09	0.23
SW	-0.30*	0.76**	-	-0.65**	0.57**	0.39**	0.36**	0.12
FE	0.35*	-0.06	-0.30*	-	-0.70**	-0.44**	-0.57**	-0.15
FS	-0.44**	-0.07	0.23	-0.72**	-	0.62**	0.69**	-0.13
FU	-0.23	0.05	0.20	-0.43**	0.51**	-	0.42**	0.10
FUHML	-0.56**	-0.01	0.36**	-0.63**	0.69**	0.33*	-	-0.40**
MV	0.27	0.30*	0.1	-0.14	-0.17	0.21	-0.35*	-

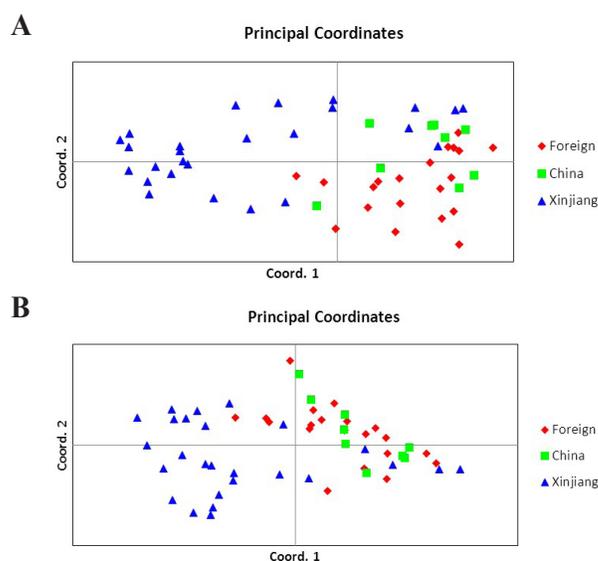
\* and \*\* = significant correlation at 0.05 and 0.01 levels, respectively. Correlation coefficients on the bottom left were the coefficients of the traits in 2009, and those on the top right were the coefficients of the traits in 2010.

### Genetic structure of 55 *G. barbadense* accessions

In order to determine the genetic structure of the 55 accessions of *G. barbadense*, separate PCA plots, based on SSR and SRAP markers, were constructed (Figure 1A and B).

Genetic structure based on SSRs revealed that the 55 *G. barbadense* accessions were divided into 3 groups (Figure 1A). The first group (accessions located left of the vertical line) predominantly consisted of Xinjiang accessions, the second group (located to the lower right of the vertical line) predominantly consisted of foreign accessions, and the third group (located to the upper right of the vertical line) was mixed.

However, the genetic structure revealed by SRAPs was slightly different from that obtained by SSRs. As shown in Figure 1B, the 55 accessions were classified into only 2 groups. The first group (located to the left of the vertical line) again consisted of mostly Xinjiang accessions, and the second group (to the right of the vertical line) was a mix of foreign and other Chinese accessions.



**Figure 1.** PCA plots of 55 *Gossypium barbadense* accessions based on SSRs (A) and SRAPs (B). Red diamonds, green squares, and blue triangles represent foreign accessions, other Chinese accessions, and Xinjiang accessions, respectively.

### Marker-trait association

A total of 170 SSRs and 258 SRAPs were used for marker-trait association after filtering for 5% minimum alleles. Based on the genotype data, the PCA matrix, the kinship matrix, and the yield and fiber quality trait data, a mixed linear model was used to analyze marker-trait associations. In order to obtain results with high confidence, we only considered significantly associated loci of  $P < 0.01$ .

There were 72 loci at the  $P < 0.01$  significance level, including 28 loci associated with yield components, and 44 loci related to fiber quality traits. The range of phenotypic variation explained (PVE) observed was from 1.35 (Me8Em10-1100) to 23.33% (NAU1164) (Tables 3 and 4).

**Table 3.** SSRs related to yield and fiber quality traits.

Trait	Year	Locus	P	PVE%	Trait	Year	Locus	P	PVE%	
LW	2009	BNL2705c	0.003	13.49	FS	2009	BNL3008	0.008	4.96	
	2010	BNL3563	0.008	15.79				TMB307b	0.004	4.03
		CIR68a	0.003	11.67			2010	TMB307b	0.008	1.76
SW	2010	BNL3563	0.008	9.11	FUHML	2009	BNL1423e	0.004	4.81	
		BNL3816a	0.007	11.59				MUCS152	0.009	5.65
		CIR68a	0.006	6.19			2010	BNL1440b	0.007	6.87
LP	2010	BNL2967b	0.000	8.34	FE		MUCS152	0.003	9.02	
		BNL3008	0.009	7.01		2009	BNL2705b	0.000	12.83	
		BNL3259b	0.006	10.46			TMB307b	0.003	7.21	
MV	2009	JESPR208b	0.008	6.02	FU	2010	BNL1017a	0.002	8.55	
		BNL1317b	0.009	10.82				BNL169b	0.007	4.99
		BNL1521c	0.006	16.22				CIR228a	0.007	7.02
		JESPR296a	0.004	2.58		2009	JESPR153b	0.005	12.23	
		TMB154a	0.003	6.27		2010	NAU1164	0.001	23.33	

PVE = phenotypic variation explained. For other abbreviations, see legend to Table 1.

**Table 4.** SRAPs related to yield and fiber quality traits.

Trait	Year	Locus	P	PVE%	Trait	Year	Locus	P	PVE%	
LW	2009	Me3Em14-120	0.008	11.28	FS	2009	Me12Em15-1300	0.000	10.95	
	2010	Me10Em14-550	0.002	14.07				Me6Em6-160	0.002	6.78
SW	2009	Me9Em4-100	0.007	14.04	MV	2010	Me9Em10-1200	0.009	4.03	
	2010	Me10Em11-2200	0.000	19.53				Me10Em11-2200	0.009	5.47
		Me10Em8-700	0.010	12.31				Me10Em14-550	0.007	4.31
		Me12Em15-1300	0.009	9.30				Me2Em13-700	0.009	3.82
		Me12Em4-550	0.004	14.45				Me9Em10-1200	0.002	5.62
		Me8Em6-100	0.007	13.15				Me1Em12-330	0.005	10.07
		Me9Em11-400	0.010	12.35				Me3Em8-260	0.000	11.80
LP	2009	Me10Em14-550	0.007	2.52	FUHML	2009	Me8Em10-700	0.007	12.36	
		Me3Em14-120	0.003	1.54				Me1Em3-270	0.000	20.06
		Me6Em17-460	0.006	4.37				Me2Em13-700	0.008	11.25
		Me9Em14-1400	0.003	5.55				Me2Em8-250	0.008	3.90
		Me10Em14-550	0.000	1.47				Me3Em8-260	0.007	2.36
		Me1Em3-270	0.010	5.00				Me6Em11-700	0.008	8.60
		Me6Em17-460	0.007	2.90				Me6Em6-160	0.004	9.71
FU	2009	Me1Em13-350	0.002	9.88	FE	2010	Me8Em7-550	0.009	5.83	
		Me11Em5-650	0.007	4.93				Me9Em10-1200	0.003	9.93
		Me3mE6-890	0.008	9.69				Me9Em2-1400	0.009	7.53
		Me6Em6-160	0.011	7.33			2009	Me2Em8-250	0.002	12.05
		Me14Em2-1600	0.010	10.60			2010	Me13Em9-850	0.002	13.70
	Me8Em10-1100	0.003	1.35		Me3Em14-120	0.007	11.16			

PVE = phenotypic variation explained. For other abbreviations, see legend to Table 1.

### Marker-trait association by SSRs

In total, 28 loci were related to yield and fiber quality traits for the 2009 and 2010 data. Ten loci were related to yield component traits, and 18 loci were related to fiber quality traits (Table 3). There were 10 loci detected in 2009, and 18 loci in 2010. Locus TMB307b was found to be related to FS in both years.

Among the 8 traits, FE was associated with the most loci, up to a maximum of 5. However, FU was only associated with 2 loci. The least phenotypic variation explained was observed in TMB307b (1.76%), which was related to FS, and the highest variation explained was in NAU1164 (23.33%), which was related to FU. The mean PVE was 8.89%, and the only traits with PVE values greater than 10% were LW and FU.

Three loci were detected for LW, one in 2009, and 2 in 2010. PVE ranged from 11.67 to 15.79%, with a mean of 13.65%. Three loci were also detected for SW, all in 2010 only, and its PVE ranged from 6.19 to 11.59%, and a mean of 8.96%. Only 4 loci were detected in 2010 for LP, with PVE range of 6.02 to 10.46%, and a mean of 7.96%. Three loci were detected for FS, including 2 loci in 2009, and 1 locus in 2010, with a PVE range of 1.76 to 4.96%, and a mean of 3.58%. For MV, 2 loci were detected in 2009 and 2 loci in 2010, with its PVE ranged from 2.58 to 16.22%, and a mean of 8.97%. Two loci were detected in 2009 and 2 loci in 2010 for FUHML, with PVE values ranging from 4.81 to 9.02%, and a mean of 6.59%. For FE, 2 loci were detected in 2009 and 3 loci were detected in 2010. PVEs for FE ranged from 4.99 to 12.83%, and mean of 8.12%. Finally, with respect to FU, 1 locus was detected in 2009 and 1 locus in 2010, with PVE range of 12.23 to 23.33%, and a mean of 17.78%.

### Marker-trait association by SRAPs

A total of 44 loci were found to be related to yield and fiber quality traits in 2009 and 2010; 16 loci were related to yield component traits and 28 loci were related to fiber quality traits (Table 4). There were 24 loci detected in 2009, and 20 loci in 2010. Three loci were detected in both 2009 and 2010: Me10Em14-550, related to LP, Me6Em17-460, also related to LP, and Me9Em10-1200, related to FS.

Among the 8 traits examined, FUHML associated with the most loci, up to 9, while FS, SW, and LP were related to 7 loci. In contrast, LW was only associated with 2 loci. The lowest observed PVE was in Me8Em10-1100 (1.35%), which was related to FU, and the highest PVE observed was in Me1Em3-270 (20.06%), which was related to FUHML. The mean PVE was 8.61%, and the traits that had PVE values greater than 10% were LW, SW, MV, and FE.

Two loci were detected for LW, 1 in 2009 and 1 in 2010, and its PVE ranged from 11.28 to 14.07%, and a mean of 12.68%. Seven loci were detected for SW, including 1 locus in 2009 and 6 loci in 2010, and its PVE ranged from 9.30 to 19.53%, and a mean of 13.59%. Seven loci were also detected for LP, including 4 loci in 2009 and 3 loci in 2010. The PVE for LP ranged from 1.47 to 5.55%, and a mean of 3.34%. Similarly, FS was associated with 7 loci, including 3 loci in 2009 and 4 loci in 2010 with PVE range of 3.82 to 10.95%, and a mean of 5.85%. Three loci were detected for MV in 2009 only, and its PVE ranged from 10.07 to 12.36%, and a mean of 11.41%. Nine loci were detected for FUHML, including 7 loci in 2009 and 2 loci in 2010, showing a PVE range of 2.36 to 20.06%, and a mean of 8.80%. Three loci were associated with FE; 1 locus was detected in 2009 and 2 loci in 2010. PVE for FE ranged from 11.16 to 13.70%, and a mean of 12.30%. Finally, 4 loci were detected for FU in 2009 and 2 loci in 2010 with a PVE range of 1.35 to 10.60%, and a mean of 7.30%.

### Comparison of association analysis by SSRs and SRAPs

The ratios of loci related to yield and fiber quality traits and respective PVE values were similar in the SSR and the SRAP analysis (Table 5). However, more traits with PVE over 10% were revealed by the SRAP analysis than by the SSR analysis. Moreover, the number of loci detected by SRAPs was close to that obtained by SSRs for both years.

**Table 5.** Comparison of association results by simple sequence repeats (SSRs) and sequence-related amplified polymorphisms (SRAPs).

	SSRs	SRAPs
Total loci	170	258
Loci related to all traits and the ratio	28 (16.47%)	44 (17.05%)
Loci related to yield and the ratio	10 (5.88%)	16 (6.20%)
Loci related to fiber quality and the ratio	18 (10.59%)	28 (10.85%)
Loci detected in 2009 and the ratio	10 (5.88%)	24 (9.30%)
Loci detected in 2010 and the ratio	18 (10.59%)	20 (7.75%)
PVE	8.89%	8.61%
Traits of PVE over 10%	2	4

PVE = phenotypic variation explained.

### DISCUSSION

False associations in QTL analysis can be caused by population structure (Flint-Garcia et al., 2003). The plant materials used in this study were 56 *G. barbadense* accessions derived from Li et al. (2008). Based on genetic distances obtained from a UPGMA analysis, we previously observed that Xinhai5 accession was distinct from the other 55 accessions, and the genetic similarity coefficient was only 0.09 (Wang et al., 2011b). Furthermore, population structure of the 56 accessions, as revealed by the program STRUCTURE, also showed that Xinhai5 had a unique genotype (data not shown). Therefore, we excluded this accession from the present study, and only used the remaining 55 accessions in genetic structure and association analysis. We here analyzed genetic structure with PCA, which showed that most of the Xinjiang accessions belonged to one group, while the other Chinese accessions and the foreign accessions clustered closely together in a separate group. Furthermore, we determined the genetic structure of these accessions using both SSR and SRAP data, and found that the boundaries among the different groups were much more defined using the SRAP dataset.

The SSR and SRAP markers used in this study were obtained from Wang et al. (2011b) and Li et al. (2008), respectively. Because the SSRs were selected from more than one linkage map, and SRAP linkage maps were rare, the chromosome positions of these SSRs and SRAPs were unknown. Therefore, the association analysis should be interpreted as marker-trait correlations.

In this study, marker-trait analysis was conducted using MLM, which considered population structure and pedigree effects as covariables in order to prevent false associations. Furthermore, the minimum allele frequency was filtered, and only significantly associated loci of  $P < 0.01$  were chosen in order to obtain associations with high confidence. The phenotypic data collected is likely a good representation of natural variation in this species since these *G. barbadense* accessions were grown in the Xinjiang region where the climate is particularly suitable for their growth. Together, these factors suggest that the association results obtained could be reliable.

TMB307b, which was related to FS, was detected both in 2009 and 2010, although the PVEs for this association were relatively low: 4.03 and 1.76%, respectively. With respect to the analysis of SRAP markers, 3 of the loci were detected in both 2009 and 2010, namely, Me10Em14-550 with LP, Me6Em17-460 with LP, and Me9Em10-1200 with FS. Further analysis showed that the ratio of loci related to yield and fiber quality traits obtained by the SSR analysis closely matched that obtained with the SRAP analysis. The PVEs of the SSR and SRAP analysis were also similar, approximately 8%, indicating that the 2 types of markers can both be effectively used for association analysis. The numbers of SRAP loci detected in 2009 and 2010 were nearly the same, suggesting that SRAPs may be less influenced by environmental fluctuations. This may be because SRAPs can amplify open reading frame regions that are less influenced by the environment than are those of SSRs.

Comparing these results to those of previous studies (Table 6), we observed 9 SSR loci that were also reported in other QTL studies, and BNL1317 was linked to MV in both this study and in Wang et al. (2011a). However, because there are fewer linkage maps based on SRAPs, none of these markers were found in previous studies.

**Table 6.** Comparison of associated SSR loci to other researches.

Our research		Other researches	
Locus	Traits	Traits	Reference
JESPR153b	FU	BN FL FL	Lin, 2005 Zhang et al., 2011 Ni et al., 2011
CIR228a	FE	Protein in seed cotton	Yu et al., 2012
JESPR208b	LP	SI LY, SY MV, FS	He et al., 2007 Luan, 2008 Wang et al., 2011a
BNL1317b	MV	MV	Wang et al., 2011a
BNL1423c,e	FUHML	SW	Lin, 2005
BNL1521c	MV	FS	Zhang et al., 2003
BNL3259b	LP	NSPB BN Oil in seed cotton Protein in seed cotton	He et al., 2007 Liu et al., 2011 Yu et al., 2012 Yu et al., 2012
BNL3563	LW, SW	SY	You, 2007

BN = boll number; NSPB = number of seeds per boll; FL = fiber length; SI = seed index; SY = seed yield; LY = lint yield. For other abbreviations, see legend to Table 1.

Both SSR and SRAP loci were linked to more than one trait, which may be caused by pleiotropism or QTL interactions. Furthermore, the PVE of these pleiotropic loci varied. For example, the PVEs of the association of Me10Em14-550 with FS, LP, and LW were 4.31, 2.52, and 14.07%, respectively, and the PVEs of Me1Em3-270 with FUHML and LP were 20.06 and 5.00%, respectively. The QTLs with high main effects and high PVEs for many traits are particularly informative. For example, the PVEs for the association of BNL3563 with LW and SW were 15.79 and 9.11%, and the PVEs of Me12Em15-1300 with LP and FS were 9.30 and 10.95%, suggesting that marker assisted selection on these traits will be particularly effective.

## ACKNOWLEDGMENTS

Research supported by the National Science Foundation of China (#30871559).

## REFERENCES

- Abdurakhmonov IY, Kohel RJ, Yu JZ, Pepper AE, et al. (2008). Molecular diversity and association mapping of fiber quality traits in exotic *G. hirsutum* L. germplasm. *Genomics* 92: 478-487.
- Abdurakhmonov IY, Saha S, Jenkins JN, Buriev ZT, et al. (2009). Linkage disequilibrium based association mapping of fiber quality traits in *G. hirsutum* L. variety germplasm. *Genetica* 136: 401-417.
- Flint-Garcia SA, Thornsberry JM and Buckler ES (2003). Structure of linkage disequilibrium in plants. *Annu. Rev. Plant Biol.* 54: 357-374.
- Grapes L, Firat MZ, Dekkers JC, Rothschild MF, et al. (2006). Optimal haplotype structure for linkage disequilibrium-based fine mapping of quantitative trait loci using identity by descent. *Genetics* 172: 1955-1965.
- He DH, Lin ZX, Zhang XL, Nie YC, et al. (2007). QTL mapping for economic traits based on a dense genetic map of cotton with PCR-based markers using the interspecific cross of *Gossypium hirsutum* x *Gossypium barbadense*. *Euphytica* 153: 181-197.
- Kantartzi SK and Stewart JM (2008). Association analysis of fibre traits in *Gossypium arboreum* accessions. *Plant Breed.* 127: 173-179.
- Li W, Ni W, Lin ZX and Zhang XL (2008). Genetic diversity analysis of sea-island cotton cultivars using SRAP markers. *Acta Agronom. Sin.* 34: 893-898.
- Lin ZX (2005). Linkage Map Construction in Cotton and QTL Mapping for Yield and Fiber-Related Traits. Doctor Dissertation of Huazhong Agricultural University, China.
- Liu R, Wang B, Guo W, Wang L, et al. (2011). Differential gene expression and associated QTL mapping for cotton yield based on a cDNA-AFLP transcriptome map in an immortalized F2. *Theor. Appl. Genet.* 123: 439-454.
- Luan MB (2008). Genetic Effect of Special Chromosomes and QTL Mapping of Main Traits. Doctor Dissertation of the Chinese Academy of Agricultural Sciences, China.
- Ni HJ, Wang W, Zhang J, Liu DJ, et al (2011). QTL Mapping of Yield and Fiber Quality Traits Using F2 and Their Derivative Populations for Upland Cotton. *Journal of Southwest University, China*, 33.
- Peakall R and Smouse PE (2006). GENALEX 6: genetic analysis in Excel population genetic software for teaching and research. *Mol. Ecol. Notes* 6: 288-295.
- Rafalski A and Morgante M (2004). Corn and humans: recombination and linkage disequilibrium in two genomes of similar size. *Trends Genet.* 20: 103-111.
- Salvi S and Tuberosa R (2005). To clone or not to clone plant QTLs: present and future challenges. *Trends Plant Sci.* 10: 297-304.
- Stich B, Melchinger AE, Frisch M, Maurer HP, et al. (2005). Linkage disequilibrium in European elite maize germplasm investigated with SSRs. *Theor. Appl. Genet.* 111: 723-730.
- Wang FR, Gong YC, Zhang CY, Liu GD, et al. (2011a). Genetic effects of introgression genomic components from Sea Island cotton (*Gossypium barbadense* L.) on fiber related traits in upland cotton (*G. hirsutum* L.). *Euphytica* 181: 41-53.
- Wang XQ, Feng CH, Lin ZX and Zhang XL (2011b). Genetic diversity of sea-island cotton (*Gossypium barbadense*) revealed by mapped SSRs. *Genet. Mol. Res.* 10: 3620-3631.
- Wen WW, Guo TT, Chavez Tovar VH, Li HH, et al. (2012). The strategy and potential utilization of temperate germplasm for tropical germplasm improvement: a case study of maize (*Zea mays* L.). *Mol. Breed.* 29: 951-962.
- Wu JX, Jenkins JN, Mc Carty JC, Zhong M, et al. (2007). AFLP marker associations with agronomic and fiber traits in cotton. *Euphytica* 153: 153-163.
- Yan WG, Li Y, Agrama HA, Luo D, et al. (2009). Association mapping of stigma and spikelet characteristics in rice (*Oryza sativa* L.). *Mol. Breed.* 24: 277-292.
- Yang XH, Yan JB, Zheng YP, Yu JM, et al. (2007). Reviews of association analysis for quantitative traits in plant. *Acta Agronom. Sin.* 33: 523-530.
- Yao J, Wang L, Liu L, Zhao C, et al. (2009). Association mapping of agronomic traits on chromosome 2A of wheat. *Genetica* 137: 67-75.
- You CY (2007). Constructon of Interspecific Linkage Map and QTL Mapping for Yield and Fiber Quality Traits Using F2 Population of *Gossypium hirsutum* x *Gossypium barbadense*. Master Dissertation of Xinjiang Agricultural University, China.
- Yu JW, Yu SX, Fan SL, Song MZ, et al. (2012). Mapping quantitative trait loci for cottonseed oil, protein and gossypol content in a *Gossypium hirsutum* x *Gossypium barbadense* backcross inbred line population. *Euphytica* 187: 191-201.
- Zhang J, Ma J, Chen X, Liu DJ, et al. (2011). QTL mapping of fiber quality traits with a composite cross population in upland cotton (*Gossypium hirsutum* L.). *J. Agr. Biotechnol.* 19: 230-235.
- Zhang T, Yuan Y, Yu J, Guo W, et al. (2003). Molecular tagging of a major QTL for fiber strength in Upland cotton and its marker-assisted selection. *Theor. Appl. Genet.* 106: 262-268.