

# QTL alleles for improved fiber quality from a wild Hawaiian cotton, *Gossypium tomentosum*

Zhengsheng Zhang · Junkang Rong · Vijay N. Waghmare ·  
Peng W. Chee · O. Lloyd May · Robert J. Wright ·  
John R. Gannaway · Andrew H. Paterson

Received: 8 April 2011 / Accepted: 22 June 2011 / Published online: 7 July 2011  
© Springer-Verlag 2011

**Abstract** Seventeen backcross-self families from crosses between two *Gossypium hirsutum* recurrent parent lines (CA3084, CA3093) and *G. tomentosum* were used to identify quantitative trait loci (QTLs) controlling fiber quality traits. A total of 28 QTLs for fiber quality traits were identified ( $P < 0.001$ ), including four for fiber elongation, eight for fiber fineness, four for fiber length, four for fiber strength, six for fiber uniformity, one for boll weight, and one for boll number. Three statistically significant marker–trait associations for lint yield were found in a single environment, but need further validation. Two-way analysis of variance revealed one locus with significant genotype  $\times$  family interaction ( $P < 0.001$ ) for fiber strength and a second locus with significant genotype  $\times$  environment

interaction ( $P < 0.001$ ) in the CA3084 background, and two loci with significant genotype  $\times$  background interaction ( $P < 0.001$ ) for the 28 common markers segregating in both of the two recurrent backgrounds. Co-location of many QTLs for fiber quality traits partially explained correlations among these traits. Some *G. tomentosum* alleles were associated with multiple favorable effects, offering the possibility of rapid genetic gain by introgression. Many *G. tomentosum* alleles were recalcitrant to homozygosity, suggesting that they might be most effectively deployed in hybrid cottons. DNA markers linked to *G. tomentosum* QTLs identified in the present study promise to assist breeders in transferring and maintaining valuable traits from this exotic source during Upland cotton cultivar development. This study also adds further evidence to prior studies indicating that the majority of genetic variation associated with fiber quality in tetraploid cotton traces to

Communicated by I. Paran.

**Electronic supplementary material** The online version of this article (doi:10.1007/s00122-011-1649-x) contains supplementary material, which is available to authorized users.

Z. Zhang · J. Rong · V. N. Waghmare · A. H. Paterson (✉)  
Plant Genome Mapping Laboratory,  
University of Georgia, Athens, GA 30602, USA  
e-mail: paterson@uga.edu

Z. Zhang  
Key Laboratory of Biotechnology and Crop Quality Improvement  
of Agricultural Ministry, Southwest University,  
Chongqing 400715, People's Republic of China

J. Rong  
College of Agricultural and Food Science,  
Zhejiang Forestry University, Hangzhou 311300,  
People's Republic of China

V. N. Waghmare  
Division of Crop Improvement,  
Central Institute for Cotton Research, Nagpur 440010, India

P. W. Chee · O. L. May  
Department of Crop and Soil Sciences,  
University of Georgia, Tifton, GA 31794, USA

R. J. Wright · J. R. Gannaway  
Texas AgriLife Research, Lubbock, TX 79403, USA

R. J. Wright  
Texas Tech University, Lubbock, TX 79409, USA

*Present Address:*  
O. L. May  
Monsanto Cotton Breeding, Tifton, GA 31793, USA

the D-subgenome from a diploid ancestor that does not produce spinnable fiber.

## Introduction

Cotton is the world's leading natural fiber crop and is an important oil and protein crop (Cherry and Leffler 1984; Lusas and Jividen 1987; Alford et al. 1996; Chen et al. 2007). The cotton genus (*Gossypium*) comprises approximately 50 species, including 45 diploids and 5 allotetraploids (Fryxell 1979; Fryxell et al. 1992; Percival et al. 1999). The allotetraploid species arose about 1–2 million years ago through the hybridization of an A-genome taxon related to the species *G. herbaceum* L. ( $2n = 2x = 26$ ), with a D-genome taxon related to the species *G. raimondii* Ulbrich and *G. gossypoides* L. ( $2n = 2x = 26$ ) (Beasley 1940, 1942; Wendel et al. 1992). The 'AD' allotetraploid ( $2n = 2x = 52$ ) group consists of the species *G. barbadense*, *G. darwinii* Watt, *G. hirsutum*, *G. tomentosum* Nuttall, and *G. mustelinum* Miers ex Watt (Percival et al. 1999).

Four *Gossypium* species are cultivated, including diploids *G. arboreum* L. and *G. herbaceum* L., and tetraploids, *G. hirsutum* L. and *G. barbadense* L. 'Upland' cotton (*G. hirsutum* L.) accounts for about 95% of the world's total production (Chen et al. 2007). However, genetic diversity among modern Upland cotton cultivars is narrow, as revealed by isozyme analysis (Wendel et al. 1992) and various DNA markers including RFLP (Brubaker et al. 1993; Brubaker and Wendel 1994), RAPD (Multani and Lyon 1995; Guo et al. 1997; Iqbal et al. 1997; Xu et al. 2001; Linos et al. 2002; Lu and Myers 2002; Zhu et al. 2003), AFLP (Abdalla et al. 2001; Iqbal et al. 2001), and SSR (Zhu et al. 2003; Rungis et al. 2005; Lacape et al. 2007). Increasing Upland cotton diversity is essential for genetic improvement. Extensive genetic variation is available among members of the genus *Gossypium* (Percival and Kohel 1990; Khan et al. 2000; Lacape et al. 2007), and efforts to increase the genetic base of *G. hirsutum* can draw upon a host of cultivated varieties, as well as primitive domesticated landraces and the other allotetraploid species.

Interspecific germplasm introgression can increase genetic variation (Percy and Wendel 1990) and has been attempted in transferring specific cotton genes and useful traits including high fiber quality (Culp et al. 1979; Cantrell and Davis 1993) and low-gossypol seeds with high-gossypol plants (Vroh Bi et al. 1998, 1999a, 1999b). However, interspecific introgression often encounters problems such as segregation distortion (Jiang et al. 2000), suppression of recombination (Paterson et al. 1990), and linkage drag (Young and Tanksley 1989), and the use of interspecific germplasm has been limited in the breeding of *G. hirsutum* and *G. barbadense*.

DNA markers provide a useful tool for detecting and resolving complications such as segregation distortion or linkage drag encountered in interspecific gene introgression (Chee et al. 2005a). In particular, the advanced backcross approach facilitates the detection and integration of beneficial quantitative trait loci (QTL) from secondary gene pools into elite breeding lines (Tanksley and Nelson 1996). By using this approach, favorable alleles for agronomically important traits have been introgressed from *G. barbadense* into *G. hirsutum* (Chee et al. 2005a, b; Draye et al. 2005; Lacape et al. 2005; Saha et al. 2004, 2006).

The objectives of the present study were to (1) identify and characterize QTLs for fiber quality traits using advanced backcross self-populations segregating for *G. tomentosum* introgressed chromatin segments, (2) analyze the interaction between *G. tomentosum* introgressed chromatin segments and genetic background, and (3) analyze the interaction between *G. tomentosum* introgressed chromatin segments and the environment.

## Materials and methods

### Population development and phenotyping

An F<sub>2</sub> population from a cross between *G. hirsutum* acc. TMS-22 and *G. tomentosum* acc. WT936 was used to develop the advanced backcross population. About 200 F<sub>2</sub> seeds were grown in Lubbock (TX) in 1998, and only about 20 plants flowered. Eleven F<sub>2</sub> plants were crossed successfully to *G. hirsutum* cultivar CA3084, and another six F<sub>2</sub> plants were crossed successfully to *G. hirsutum* cultivar CA3093. From 1999 to 2001, the 17 F<sub>2</sub>BC<sub>1</sub> families were backcrossed to CA3084 or CA3093, respectively, obtaining 17 BC<sub>3</sub>F<sub>1</sub> families. Seeds of all BC<sub>3</sub>F<sub>1</sub> families were planted in peat pellets and germinated in the greenhouse, and seedlings were hand-planted in the field (Lubbock, TX) during early April of 2002. Plants were spaced 30 cm apart within rows and rows were 152 cm apart. A total of 634 BC<sub>3</sub>F<sub>2</sub> plants from two background populations were used to extract DNA and also were self-pollinated to produce BC<sub>3</sub>F<sub>3</sub> populations in 2002. Eleven BC<sub>3</sub>F<sub>2</sub> families in the CA3084 background included 319 plants ranging from 2 to 94 plants per BC<sub>3</sub>F<sub>1</sub>-derived family, and six BC<sub>3</sub>F<sub>2</sub> families in the CA3093 background included 315 plants ranging from 16 to 93 lines per BC<sub>3</sub>F<sub>1</sub>-derived family. In 2003, the BC<sub>3</sub>F<sub>3</sub> populations were grown in the field in Lubbock (TX) and Tifton (GA), respectively. Except for spacing (which was as described for the BC<sub>3</sub>F<sub>2</sub> plants), cultural practices were normal for cotton production in Texas and Georgia. Seed cotton from mature bolls of BC<sub>3</sub>F<sub>2</sub> and BC<sub>3</sub>F<sub>3</sub> plants were hand-harvested and ginned on a saw gin. Some plants were sterile or produced insufficient lint for

fiber analysis. Fiber samples from 218 progenies in the CA3084 background (ranging from 1 to 75 plants per family), and 243 progenies in the CA3093 background (16–73 plants per family) were sent to the Cotton Incorporated Textile Services Laboratory (Cotton Incorporated, Cary, NC) where fiber quality traits were determined by a High-Volume Precision Instrument (HVI; Zellweger-Uster, Knoxville, Tenn.). Lint percentage and lint yield of BC<sub>3</sub>F<sub>3</sub> populations were only determined in Tifton (GA) in 2003.

#### Genotyping and data analysis

The genome composition of the introgression populations was obtained by genotyping BC<sub>3</sub>F<sub>2</sub> plants with 448 RFLP markers from a previously constructed genetic map (TH map) containing 589 loci and covering 4,259.4 cM (Waghmare et al. 2005). A total of 97 and 62 informative RFLP markers were identified and genotyped for CA3084 and CA3093 background populations, respectively.

Associations between marker genotypes and fiber quality and/or yield traits were tested for statistical significance by one-way variance analyses for each marker locus segregating in the BC<sub>3</sub>F<sub>2</sub>/BC<sub>3</sub>F<sub>3</sub> populations, using the GLM procedure of SAS 9.2 (SAS Institute Inc. 2008) and an *F* test significance threshold of  $P < 0.001$ . The  $R^2$  of each variance analysis provided an estimate of the proportion of phenotypic variance explained by the corresponding marker locus. For significant marker–trait associations, the phenotypic effect of individual QTLs was estimated by MapQTL 6.0 (Van Ooijen 2009). The extreme paucity of homozygotes made it impossible to estimate gene action (additivity and dominance) at most loci, so the analysis of QTL effects in this population was essentially limited to estimating phenotypic effects of a single allele substitution. The QTL map was presented by MapChart 2.2 (Voorrips 2006). QTL nomenclature follows a method used in rice (McCouch et al. 1997), starting with ‘q’, followed by an abbreviation of the trait name [fiber elongation (FE), fiber fineness (FF) (Micronaire), fiber length (FL), fiber strength (FS), fiber length uniformity (FU), boll weight (BW), 25 boll numbers (BN), lint percent (LP), Lint yield (LY)] and the name of chromosome, then followed by a number. While most cotton fiber traits have intuitive definitions, we note that ‘fiber elongation’ refers to the maximum extension at which the yarn (composed of individual fibers) breaks.

For loci segregating in two or more families of one background, or in both backgrounds, two-way mixed model variance analyses were performed using the MIXED procedure of SAS 9.2 (SAS Institute Inc. 2008). The analytical model included genotype (G), family (F) and genotype  $\times$  family ( $G \times F$ ) interaction, or background (B) and genotype  $\times$  background ( $G \times B$ ) interaction as

fixed factors. The interaction between genotype (G) and environment (E) was also analyzed by the MIXED procedure. Model parameters were estimated using the residual maximum likelihood (REML) method. Marker–trait association (genotype factor) was tested with an *F* statistic using a general Satterthwaite approximation for the denominator degrees of freedom (SAS Institute Inc. 2008). A likelihood-ratio statistic (ChiSq) was performed for the  $G \times F$ ,  $G \times B$  and  $G \times E$  interaction (Self and Liang 1987). Genotypic effects and interaction effects were considered significant if  $P < 0.001$ . This stringent threshold, similar to an LOD score of 3.0, is used to achieve an experiment-wise false positive rate no higher than 0.05.

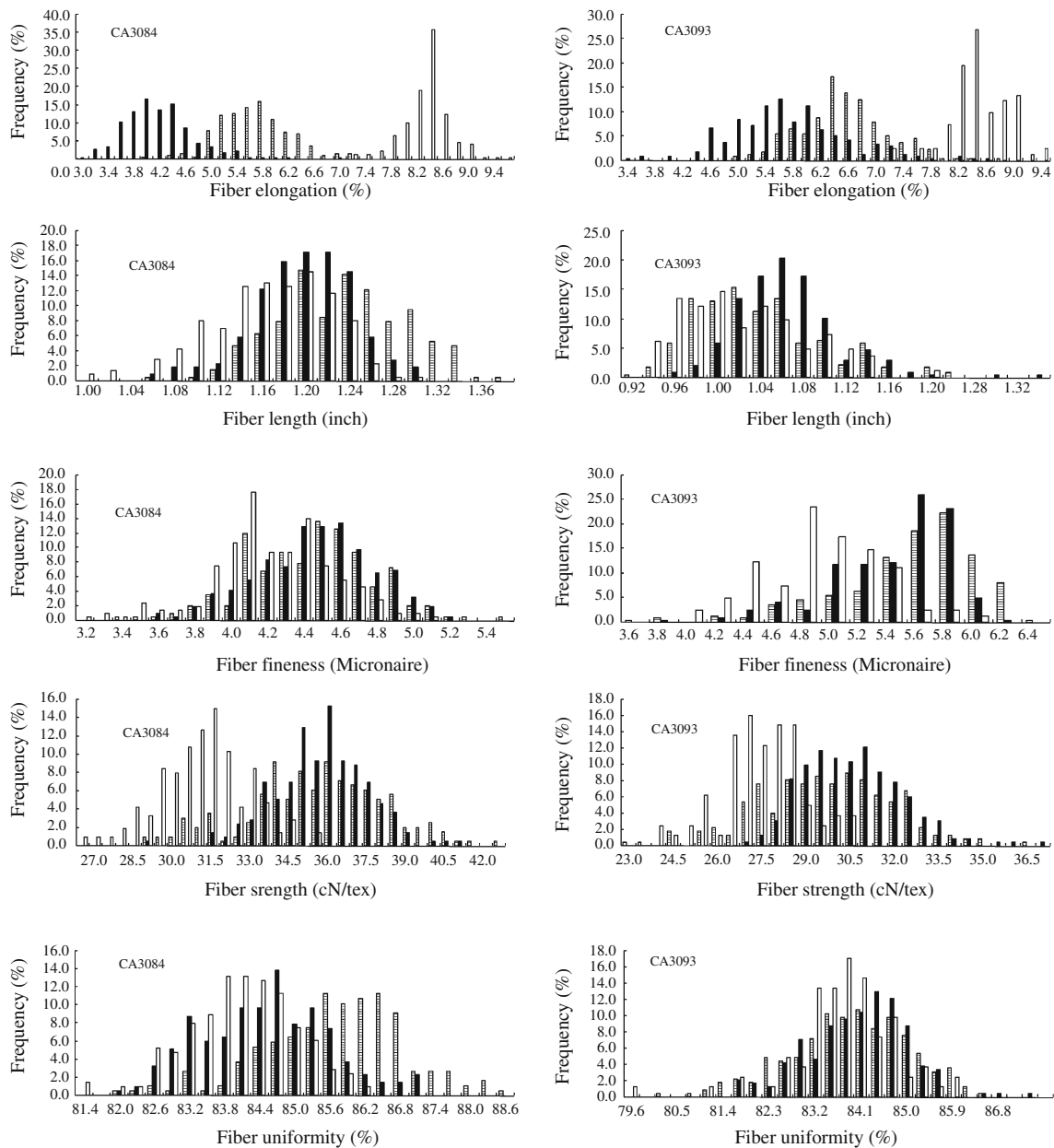
## Results

### Population structure

Out of 448 informative RFLP markers in the TH genetic map, 129 showed segregation in at least one of the two background populations. Among these markers, 97 (average 8.8 per family) segregated among 11 BC<sub>3</sub>F<sub>1</sub> families in the CA3084 background population, whereas 62 (average 10.3 per family) segregated among six BC<sub>3</sub>F<sub>1</sub> families in the CA3093 background population. The 17 BC<sub>3</sub>F<sub>1</sub> families included 461 BC<sub>3</sub>F<sub>2</sub> individuals with phenotypic data (218 for CA3084 background, 243 for CA3093) ranging from 1 to 75 plants per family. The average allele introgression ratio is 4.32% (4.49% in CA3084, 4.06% in CA3093), much lower than the Mendelian expectation (12.5%). The introgressed chromosome segments covered a total of 1,010.0 cM (719.7 cM in CA3084, 546.3 cM in CA3093, 256 cM overlapping) and accounted for 23.7% (16.9% in CA3084, 12.8% in CA3093) of the TH genetic map. The total length of introgressed *G. tomentosum* segments per chromosome ranged from 7.9 cM (Chr. 2) to 131.6 cM (Chr. 19) with an average of 47.8 cM. A total of 28 overlapping introgressions were found in the two backgrounds.

### Phenotypic performance

The statistical parameters reflecting fiber quality and yield traits are listed in Suppl Table s1 and the distributions of traits are presented in Fig. 1. The fiber quality and yield traits are normally distributed except for lint percentage of the CA3093 population and lint yield of both populations. For fiber length, fiber length uniformity, fiber strength and lint yield, trait means of the CA3084 background population were substantially and significantly higher ( $P < 0.0001$ ) than those of the CA3093 background population, except that the higher fiber length uniformity of the CA3084 BC<sub>3</sub>F<sub>3</sub> population in one location (Tifton, GA)



**Fig. 1** Frequency distribution of fiber quality and yield traits in the  $BC_3F_2/BC_3F_3$  populations. Plain, hashed and empty bars indicate  $BC_3F_2$  at Lubbock (TX),  $BC_3F_3$  at Lubbock (TX), and  $BC_3F_3$  at Tifton (GA), respectively

was not statistically significant. Fiber fineness (Micronaire) is significantly lower for the CA3084 background population than the CA3093 population, another desirable attribute of CA3084. However, for fiber elongation and lint percentage, the trait means of the CA3093 background population are substantially and significantly higher ( $P < 0.0001$ ) than those of the CA3084 population, except that the advantage of CA3093 for fiber elongation of the  $BC_3F_3$  in one location (Tifton, GA) only reaches  $P < 0.005$ .

Analyses of variance of fiber quality traits are presented in Table 1. Fiber quality traits showed highly significant

genetic and environment effects ( $P < 0.0001$ ), except that fiber uniformity of the CA3093 background population had a less significant genetic effect ( $P < 0.001$ ) and non-significant environment effect. Among the five fiber quality traits, fiber elongation is the most affected by environment. For example, the minimum value in the CA3084  $BC_3F_3$  background population in one environment (Tifton, GA) is larger than the maximum in the other environment (Lubbock, TX) (Table S1; Fig. 1).

Correlations among the five fiber quality traits generally followed patterns that are characteristic of cotton germplasm

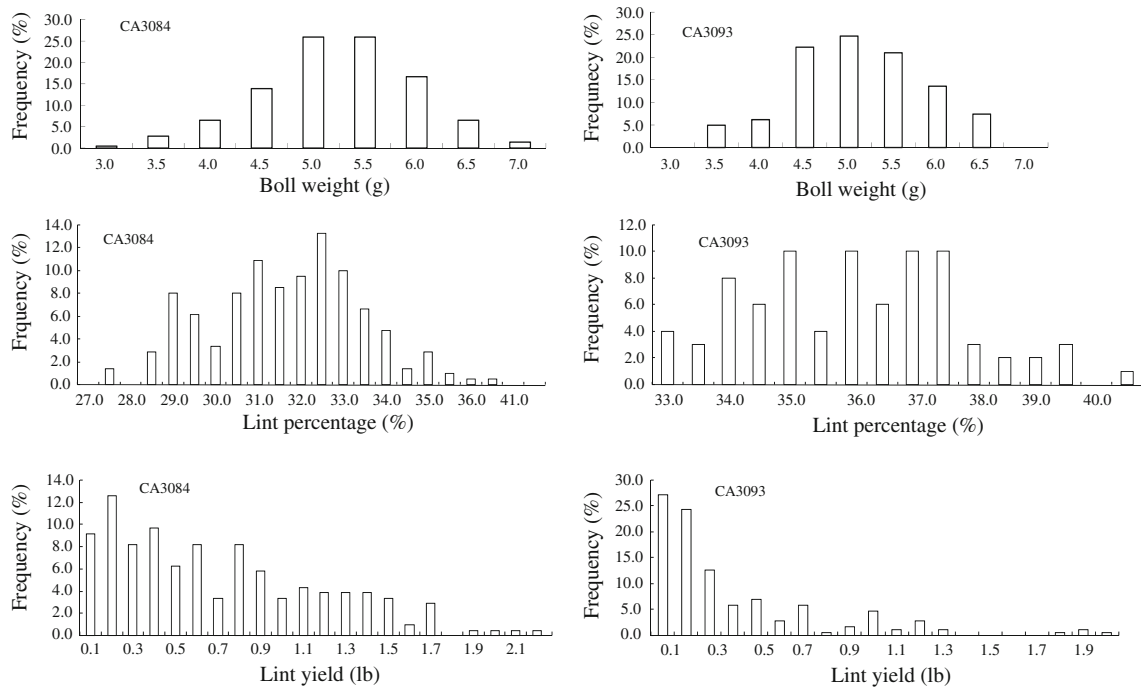


Fig. 1 continued

**Table 1** Variance analysis of fiber quality and yield traits in two background populations

Trait	Source of variation	DF		MS		F	
		CA3084	CA3093	CA3084	CA3093	CA3084	CA3093
Elongation	Genotype	217	242	0.5796	1.1555	6.18**	65.25**
	Environment	2	2	892.2597	219.9906	9,518.33**	12,422.6**
	Error	402	301	0.09374	0.01771		
Fineness (Micronaire)	Genotype	217	242	0.2087	0.3514	3.5**	4.49**
	Environment	2	2	3.8854	11.6916	65.12**	149.43**
	Error	402	301	0.0597	0.0782		
Length	Genotype	217	242	0.0062	0.0051	6.02**	3.73**
	Environment	2	2	0.2335	0.0536	227.82**	39.12**
	Error	402	301	0.001	0.0014		
Strength	Genotype	217	242	6.8041	5.979	2.57**	3.23**
	Environment	2	2	1,464.724	273.4485	552.75**	147.93**
	Error	402	301	2.6499	1.8485		
Uniformity	Genotype	217	242	1.7764	1.3682	2.21**	1.54*
	Environment	2	2	145.1317	2.2468	180.38**	2.53
	Error	402	301	0.8046	0.8884		

\* and \*\* represent significance with  $P$  values of 0.01 and 0.001, respectively

(Table 2). Fiber length, strength, and uniformity are significantly positively related to one another and significantly negatively related to fiber fineness (where lower values represent higher quality) and elongation. High lint percentage was weakly associated ( $r = 0.18$ ,  $p < 0.01$ ) with increased (undesirable) micronaire in CA3084, and strongly associated ( $r = -0.63$ ,  $p < 0.001$ ) with reduced fiber length in

CA3093. Boll weight was strongly associated with reduced lint yield in both CA3084 and CA3093. High lint yield was weakly associated with both increased fiber length ( $r = 0.23$ ,  $p < 0.001$ ) and increased lint percentage ( $r = 0.20$ ,  $p < 0.01$ ) in CA3084.

Correlations of fiber quality and yield traits across years and environments are generally high (Table 3), with only



**Table 2** Correlation coefficients between different fiber quality and yield traits

Trait	Environment	Elongation	Micronaire	Length	Strength	Uniformity	Boll weight	Lint percentage
Fiber fineness (Micronaire)	BC3F2 (TX)	0.13/0.18*						
	BC3F3 (GA)	0.16/0.47**						
	BC3F3 (TX)	0.16/0.10						
Fiber length	BC3F2 (TX)	-0.49**/-0.52**	-0.37**/-0.61**					
	BC3F3 (GA)	-0.36**/-0.48**	-0.21*/-0.38**					
	BC3F3 (TX)	-0.37**/-0.38**	-0.46**/-0.69**					
Fiber strength	BC3F2 (TX)	-0.19*/-0.26**	-0.06/-0.03	0.26**/0.23**				
	BC3F3 (GA)	0.23**/0.34*	-0.16/0.20	0.18*/0.15				
	BC3F3 (TX)	-0.27**/-0.25**	-0.17/-0.13	0.21*/0.38**				
Fiber uniformity	BC3F2 (TX)	-0.01/-0.08	-0.03/0.21*	0.38**/0.20*	0.55**/0.49**			
	BC3F3 (GA)	0.33**/0.34*	0.06/0.23	0.21*/-0.13	0.22*/0.18			
	BC3F3 (TX)	0.11/0.07	-0.10/0.25**	0.30**/0.08	0.32**/0.34**			
Boll weight	BC3F3 (GA)	-0.06/0.27*	0.29**/0.23	0.28**/0.07	0.01/0.21	0.06/0.30*		
Lint percentage	BC3F3 (GA)	0.04/0.26	0.18*/0.23	-0.07/-0.63**	-0.06/-0.21	-0.12/0.01	0.21*/-0.03	
Lint yield	BC3F3 (GA)	0.05/0.14	-0.01/0.14	0.23**/0.12	-0.11/0.14	0.08/0.23	0.34**/0.46**	0.20*/-0.15

The numerical value on the left of the slash is for CA3084 background and on the right is for CA3093 background

BC<sub>3</sub>F<sub>2</sub> (TX), BC<sub>3</sub>F<sub>3</sub> (TX) and BC<sub>3</sub>F<sub>3</sub> (TX) indicate BC<sub>3</sub>F<sub>2</sub> at Lubbock (TX), BC<sub>3</sub>F<sub>3</sub> at Lubbock (TX), and BC<sub>3</sub>F<sub>3</sub> at Tifton (GA), respectively

\* and \*\* represent significance with *P* values of 0.01 and 0.001, respectively

**Table 3** Correlation coefficients of fiber quality traits across years and environments

Trait	BC <sub>3</sub> F <sub>2</sub> (TX) versus BC <sub>3</sub> F <sub>3</sub> (TX)	BC <sub>3</sub> F <sub>2</sub> (TX) versus BC <sub>3</sub> F <sub>3</sub> (GA)	BC <sub>3</sub> F <sub>3</sub> (TX) versus BC <sub>3</sub> F <sub>3</sub> (GA)
Elongation	0.62**/0.67**	0.47**/0.37**	0.53**/0.41**
Length	0.66**/0.60**	0.57**/0.61**	0.69**/0.38**
Micronaire	0.47**/0.52**	0.40**/0.38**	0.52**/0.34*
Strength	0.40**/0.28**	0.27**/0.46**	0.38**/0.35*
Uniformity	0.36**/0.24**	0.17/0.24	0.30**/0.13

The numerical value on the left of the slash is for the CA3084 background and on the right is for CA3093 background

BC<sub>3</sub>F<sub>2</sub> (TX), BC<sub>3</sub>F<sub>3</sub> (TX) and BC<sub>3</sub>F<sub>3</sub> (TX) indicate BC<sub>3</sub>F<sub>2</sub> at Lubbock (TX), BC<sub>3</sub>F<sub>3</sub> at Lubbock (TX), and BC<sub>3</sub>F<sub>3</sub> at Tifton (GA), respectively

\* and \*\* represent significance at *P* values of 0.01 and 0.001, respectively

fiber length uniformity failing to be statistically significant between virtually all years and backgrounds. This result showed that fiber quality traits had relatively high heritability, except for fiber length uniformity.

### Main-effect QTL

A total of 3,180 marker–trait associations (1,940 in the CA3084 background population, 1,240 in CA3093) were tested, and 73 reached statistical significance. These marker–trait associations are presented in a supplemental document (Suppl Table s2). Thirty-one significant marker–trait associations were on 11 A-genome chromosomes and 42 were on 11 D-genome chromosomes.

However, segregation distortion was extreme in this population, with some loci containing only one *G. tomentosum* allele in the entire population. While it is possible to statistically exclude the phenotype of one individual from being part of the same population as the remaining segregants (and all such statistically significant comparisons are shown in Suppl Table s2), it is not possible to ascribe a phenotypic effect to the unusual allele because a single individual is clearly insufficient to provide a sample that is ‘homogenized’ for the remainder of the genome. For example, we found statistically significant associations of fiber elongation with pGH276 on Chr. 6, and pAR873 on Chr. 9 that were due to the presence of *G. tomentosum* alleles at these two loci in the same individual. Further data would be necessary to distinguish which of these loci, or perhaps others, are responsible for the phenotypic difference. Among those codominant loci that remain heterozygous in BC<sub>3</sub>F<sub>1</sub>, heterozygotes occurred at an average frequency of 28.9% whereas *G. tomentosum* homozygotes occurred at 1.8% in BC<sub>3</sub>F<sub>2</sub>—in other words, only about one individual in three harbor at least one *G. tomentosum* allele at a locus that remains heterozygous in the BC<sub>3</sub>F<sub>1</sub>. To be 99% confident that individuals that all share one or more *G. tomentosum* alleles at one locus do not also all share *G. tomentosum* alleles at another unlinked locus, about 15 individuals must be examined. Accordingly, while we have listed all statistically significant marker–trait associations in the supplemental document (Suppl Table s1), we have only considered those cases in which the smaller genotypic class included at least 15 individuals as representing QTLs (Table 4).

**Table 4** QTL for fiber-related traits in *G. tomentosum* chromosome introgressed populations

QTL	Chromosome	Nearest marker	HH:HT:TT or HH:HT or (T-)	Background	Environment	CA3084 background		CA3093 background		g × b
						Effect	PVE%	Effect	PVE%	
Fiber elongation (%)										
qFE11.1	11	pBAM422yE3C	207:15	CA3093	BC <sub>3</sub> F <sub>2</sub> (TX)			−0.66**	6.7	
qFE14.1	14	pAR815E3C	199:23	CA3084, CA3093	BC <sub>3</sub> F <sub>2</sub> (TX)			−0.64**	9.0	
			206:32	CA3084, CA3093	BC <sub>3</sub> F <sub>3</sub> (TX)			−0.53*	4.5	
qFE19.1	19	pAR847xE3R	174:(43)	CA3084	BC <sub>3</sub> F <sub>3</sub> (TX)	−0.27**	8.9			
			174:42	CA3084	BC <sub>3</sub> F <sub>3</sub> (GA)	−0.14*	5.6			
qFE21.1	21	G1261aE3C	204:15	CA3093	BC <sub>3</sub> F <sub>2</sub> (TX)			−0.72**	7.8	
			215:20	CA3093	BC <sub>3</sub> F <sub>3</sub> (TX)			−0.72*	5.7	
Fiber fineness (micronaire)										
qFF04.1	04	pAR3-46yE3C	199:16	CA3084	BC <sub>3</sub> F <sub>3</sub> (GA)	0.31*	6.3			
qFF05.1	05	pAR1-28E3C	195:20	CA3084, CA3093	BC <sub>3</sub> F <sub>3</sub> (GA)	0.32**	8.1			
qFF07.1	07	G1158bE5C	212:17	CA3093	BC <sub>3</sub> F <sub>3</sub> (TX)			0.40**	6.0	
qFF13.1	13	A1135yE3R	208:(26)	CA3093	BC <sub>3</sub> F <sub>3</sub> (TX)			0.18*	5.0	
qFF14.1	14	pAR815E3C	206:32	CA3084, CA3093	BC <sub>3</sub> F <sub>3</sub> (TX)			0.42**	11.9	**
qFF15.1	15	pAR264E5C	208:22	CA3093	BC <sub>3</sub> F <sub>3</sub> (TX)			0.41**	8.1	
qFF19.1	19	pAR282E3R	163:(53)	CA3084	BC <sub>3</sub> F <sub>3</sub> (TX)	−0.11*	5.5			
qFF21.1	21	G1261aE3C	215:20	CA3093	BC <sub>3</sub> F <sub>3</sub> (TX)			0.33*	5.1	
Fiber length (inches)										
qFL04.1	04	pAR3-46yE3C	201:16	CA3084	BC <sub>3</sub> F <sub>3</sub> (TX)	0.04**	6.3			
			199:16	CA3084	BC <sub>3</sub> F <sub>3</sub> (GA)	0.05**	6.5			
qFL08.1	08	A1401zE5C	179:19	CA3084	BC <sub>3</sub> F <sub>3</sub> (GA)	0.05**	7.0			
qFL08.2	08	pAR3-7E4C	179:19	CA3084	BC <sub>3</sub> F <sub>3</sub> (GA)	0.05**	7.0			
qFL15.1	15	pGH317zE4R	192:(22)	CA3084	BC <sub>3</sub> F <sub>3</sub> (GA)	0.03**	8.3			
Fiber strength (cN/tex)										
qFS02.1	02	pAR316E4C	153:82:1	CA3093	BC <sub>3</sub> F <sub>3</sub> (TX)			−1.13**	12.9	
			54:27	CA3093	BC <sub>3</sub> F <sub>3</sub> (GA)			−1.08*	13.1	
qFS15.1e	15	A1720xE4R	104:(73)	CA3084	BC <sub>3</sub> F <sub>2</sub> (TX)	−0.89**	8.3			
qFS16.1	16	pAR285yE3C	191:49	CA3093	BC <sub>3</sub> F <sub>3</sub> (TX)			0.89*	4.7	
qFS22.1	22	pAR949xE4C	145:50:4	CA3084	BC <sub>3</sub> F <sub>3</sub> (GA)	−0.75*	6.1			
Fiber uniformity (%)										
qFU14.1	14	pAR815E3C	206:32	CA3093	BC <sub>3</sub> F <sub>3</sub> (TX)			0.63*	5.4	**
qFU15.1	15	A1720xE4R	104:(73)	CA3084	BC <sub>3</sub> F <sub>2</sub> (TX)	−0.42*	8.3			
qFU15.2	15	pAR264E5C	208:22	CA3093	BC <sub>3</sub> F <sub>3</sub> (TX)			0.76*	5.9	
qFU20.1	20	pAR956E5C	154:52	CA3084	BC <sub>3</sub> F <sub>3</sub> (TX)	−0.61*	6.5			
qFU22.1	22	pAR949xE4C	145:50:4	CA3084	BC <sub>3</sub> F <sub>3</sub> (GA)	−0.47*	6.6			
qFU26.1	26	pGH413E4C	143:20:11	CA3084	BC <sub>3</sub> F <sub>2</sub> (TX)	−0.53*	5.7			
Boll number (25 bolls)										
qBN21.1	21	pAR944E3C	199:34:5	CA3093	BC <sub>3</sub> F <sub>3</sub> (GA)			*	6.2	
Boll weight (g)										
qBW04.1	04	pAR3-46yE3C	200:16	CA3084	BC <sub>3</sub> F <sub>3</sub> (GA)	0.87**	9.9			
Lint yield (lb)										
	15	pGH317zE4R	192:(22)	CA3084	BC <sub>3</sub> F <sub>3</sub> (GA)	0.23*	5.2			
	20	pAR956E5C	154:55	CA3084	BC <sub>3</sub> F <sub>3</sub> (GA)	−0.25*	5.4			
	22	pAR949xE4C	145:50:4	CA3084	BC <sub>3</sub> F <sub>3</sub> (GA)	−0.28*	8.9			

BC<sub>3</sub>F<sub>2</sub> (TX), BC<sub>3</sub>F<sub>3</sub> (TX) and BC<sub>3</sub>F<sub>3</sub> (TX) indicate BC<sub>3</sub>F<sub>2</sub> at Lubbock (TX), BC<sub>3</sub>F<sub>3</sub> at Lubbock (TX), and BC<sub>3</sub>F<sub>3</sub> at Tifton (GA), respectively  
PVE phenotypic variance explained

F and E indicate significant interaction ( $P < 0.001$ ) for genotype × family lines and genotype × environment, respectively

+ indicates that the *G. hirsutum* allele increases the trait value, and − indicates that the *G. tomentosum* allele increases the trait value

\* and \*\* represent significance with  $P$  values of 0.001 and 0.0001, respectively

### Fiber elongation

Four QTLs were identified, including one in the CA3084 background population and three in CA3093 (Table 4; Fig. 2). Two of the four QTLs were identified in each of 2 years, and one in each of two locations. The phenotypic effects of these QTLs ranged from  $-0.72$  to  $-0.14$ , explaining from 4.5 to 8.9% of phenotypic variance. The favorable alleles of all four QTLs originated from *G. tomentosum*. One QTL was located on an A-subgenome chromosome (11) and three on D-subgenome chromosomes (14, 19, 21).

### Fiber fineness

Eight QTLs were identified in the BC<sub>3</sub>F<sub>3</sub> generation (none in BC<sub>3</sub>F<sub>2</sub>), two in GA and six in TX, including three in CA3084 background and five in CA3093 (Table 4; Fig. 2). The phenotypic effects of these QTLs ranged from  $-0.11$  to 0.42 micronaire units, explaining from 5.1 to 11.9% of phenotypic variance. The favorable alleles of seven QTLs (87.5%) originated from *G. tomentosum* and one (12.5%) from *G. hirsutum*. Four QTLs were located on A-subgenome chromosomes and four on D-subgenome chromosomes.

### Fiber length

Four QTLs were identified on three chromosomes in the CA3084 background (Table 4; Fig. 2). The phenotypic effects of these QTLs ranged from 0.03 to 0.05 inches, explaining from 6.3 to 8.3% of phenotypic variance. One QTL (FL04.1) was identified in two environments. The favorable alleles of all four QTLs originated from *G. hirsutum*. Three QTLs were located on A-subgenome chromosomes and one on a D-subgenome chromosome.

### Fiber strength

Four QTLs were identified, including two in the CA3084 background and two in CA3093 (Table 4; Fig. 2). One (qFS02.1) was identified in two environments. The phenotypic effects of these QTLs ranged from  $-1.13$  to 0.89 cN/tex, explaining from 4.7 to 13.1% of phenotypic variance. The favorable alleles of three QTLs originated from *G. tomentosum* and one from *G. hirsutum*. One QTL was located on an A-subgenome chromosome and three QTLs were on D-subgenome chromosomes. One additional QTL, qFS05.1 shows no main effect but significant interaction between family lines.

### Fiber uniformity

Six QTLs were identified on five chromosomes, including four in CA3084 background population and two in

CA3093 (Table 4; Fig. 2). Each QTL was detected in only one environment. The phenotypic effects of these QTLs ranged from  $-0.61$  to  $+0.76$ , explaining from 5.4 to 8.3% of phenotypic variance. The favorable alleles of four QTLs originated from *G. tomentosum* and two from *G. hirsutum*. All QTLs were located on D-subgenome chromosomes.

### Boll number

One QTL was identified on Chr. 21 of the D-subgenome in the CA3093 background (Table 4; Fig. 2) in the BC<sub>3</sub>F<sub>3</sub> generation only. This QTL explained 6.2% of phenotypic variance. The favorable allele originated from *G. hirsutum*.

### Weight per boll

One QTL was identified on Chr. 4 of the A-subgenome in the CA3084 background (Table 4; Fig. 2). The phenotypic effect was 0.87, explaining 9.9% of phenotypic variance. The favorable allele originated from *G. hirsutum*.

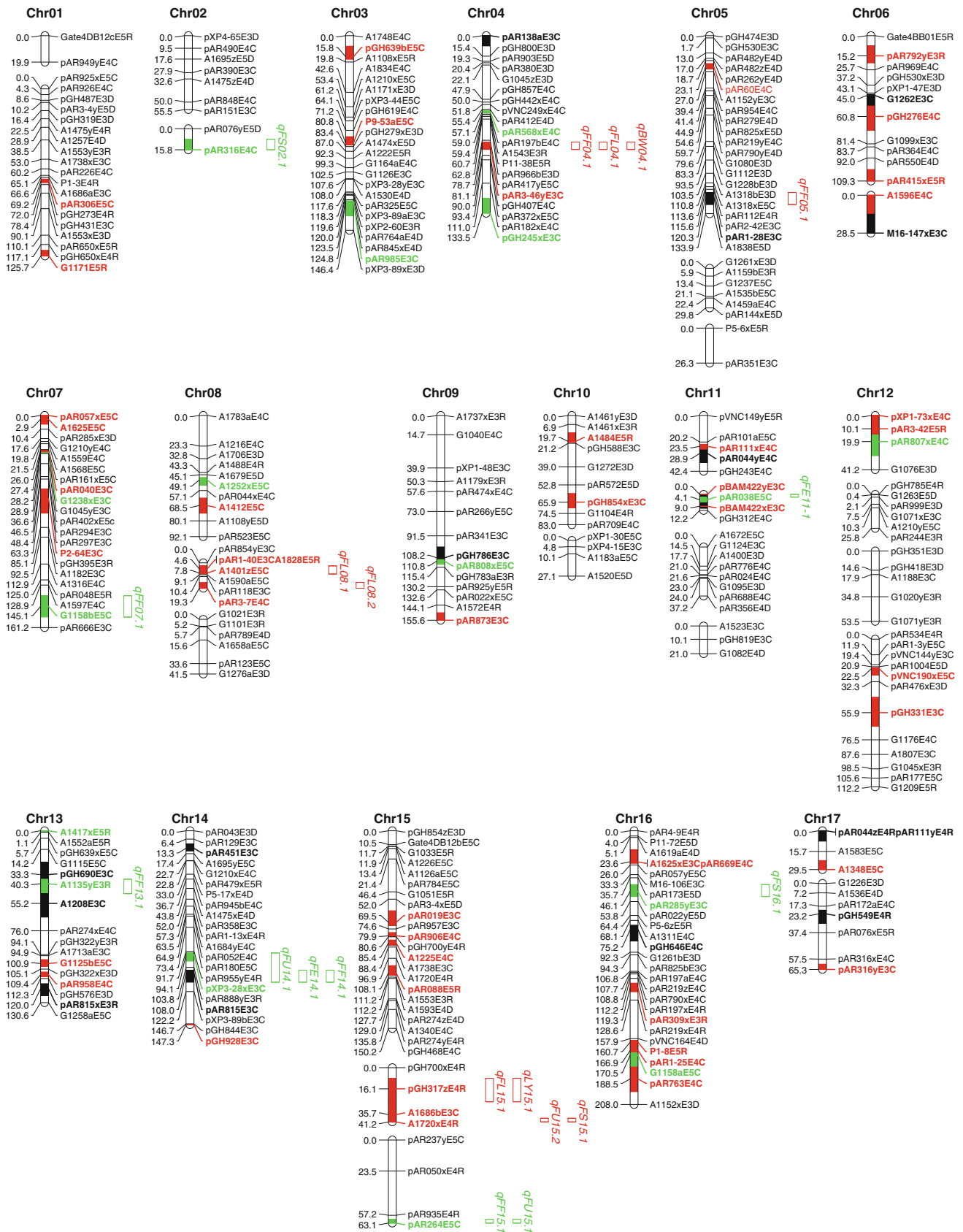
### Lint yield

Three statistically significant marker–trait associations were identified on three D-genome chromosomes in CA3084 background population (Table 4; Fig. 2). The phenotypic effects of the QTLs ranged from  $-0.28$  to 0.23, explaining from 5.2 to 8.9% of phenotypic variance. The favorable alleles at two loci originated from *G. tomentosum* and one from *G. hirsutum*. Because lint yield is a complex trait of low heritability and we have data from only a single environment (GA BC<sub>3</sub>F<sub>3</sub>), we have not formally claimed these associations to be QTLs, but note them here to facilitate future meta-analyses that may provide corroborative data.

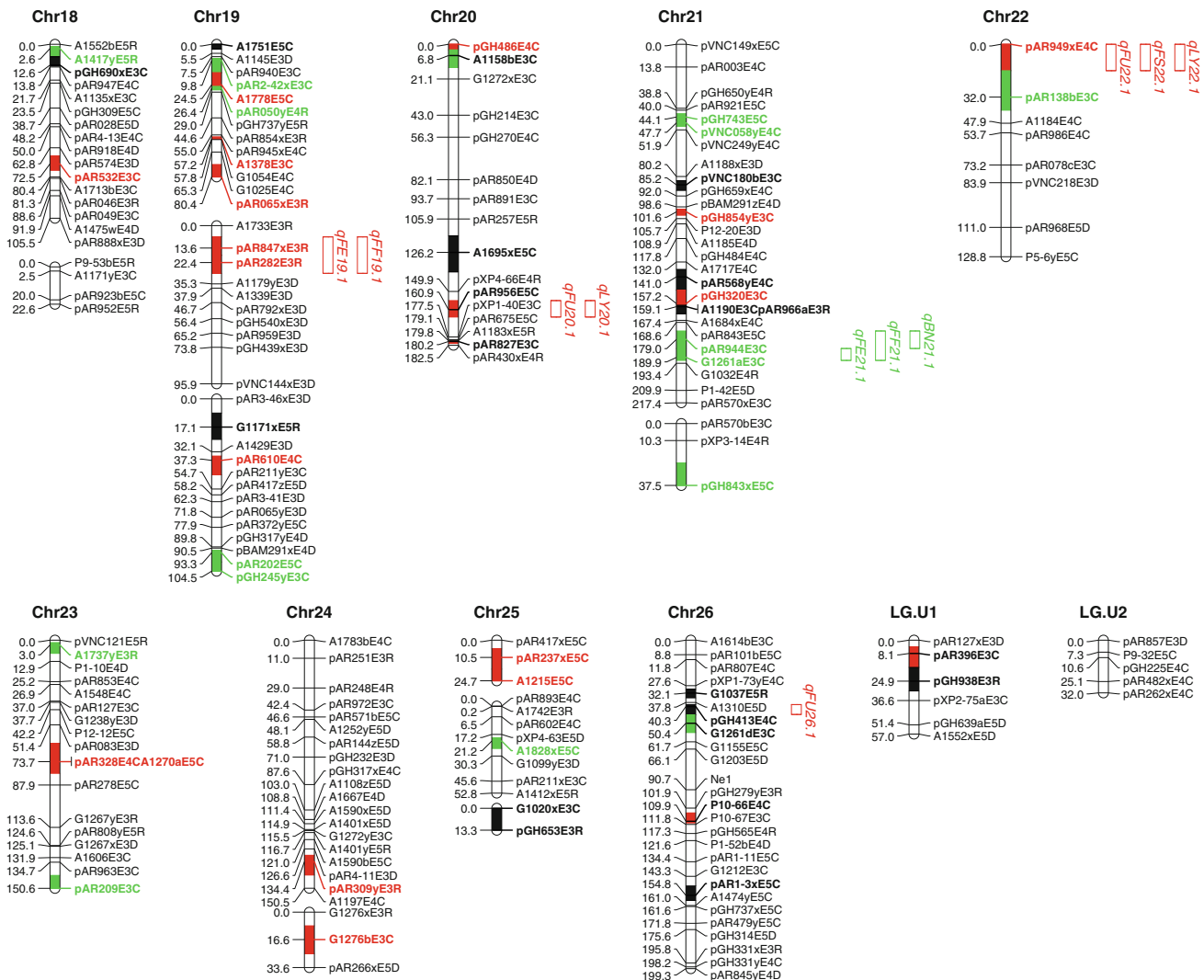
### Interaction between genotype and background

Only one genotype  $\times$  family interaction was found for the 39 markers segregating in two or more family lines in the CA3084 background population, whereas no genotype  $\times$  family interaction was found for the 25 markers segregating in two or more family lines in the CA3093 background population. However, for the significant genotype  $\times$  family interaction locus, FS05.1, the QTL main effect, was not significant. For the 28 markers segregating in both CA3084 and CA3093 background populations, seven significant genotype  $\times$  background interactions were found but only two, qFF14.1 and qFU14.1, each diagnosed by the same RFLP marker (pAR815E3C) (Table 4), showed significant main effects.





**Fig. 2** QTLs for fiber quality and yield traits identified in *G. tomentosum* introgression Upland cotton populations. Red, green, and black bars indicate the *G. tomentosum* chromosome segments in CA3084, CA3093, and both CA3084 and CA3093 background populations, respectively



**Fig. 2** continued

## Interaction between genotype and environment

For the five fiber quality traits, a total of 97 marker–environment associations were tested in the CA3084 background population, and only one genotype × environment interaction was found, involving qFS15.1 (Table 4). No interaction was found among the 62 marker–environment combinations in the CA3093 background population.

## Discussion

### Favorable QTL alleles from *G. tomentosum*, a wild cotton

Among the QTLs identified, favorable alleles originated from *G. tomentosum* for 18 (64.3%) and *G. hirsutum* for 10 (35.7%). *G. hirsutum* contributed all four favorable alleles for fiber length (all in the CA3084 background),

and *G. tomentosum* contributed most favorable alleles for fiber elongation and fiber fineness. This result is consistent with the finding that favorable QTL alleles may be recovered from an apparently unfavorable parent (Tanksley and Nelson 1996; Xiao et al. 1996), of particular importance to cotton in view of the series of genetic bottlenecks that have constrained variation in its present gene pool.

It may be noteworthy that the traits, for which *G. tomentosum* contributed the preponderance of favorable alleles, fiber fineness and elongation, have a relatively short history of selection in scientific cotton breeding. Differences between the CA3084 and 3093 backgrounds show that there is heterogeneity for favorable traits in the *G. hirsutum* gene pool. For traits that have been under selection in *G. hirsutum* for a long time, repeatedly crossing the best with the best would ‘concentrate’ favorable alleles. The cases in which *G. tomento-*

*sum* conferred most of the favorable alleles may suggest that (for reasons we do not understand) the favorable alleles for those traits are at higher frequency in *G. tomentosum* than *G. hirsutum*. So, it may be more likely that traits for which *G. tomentosum* had many favorable alleles are those that had not been under long or strong selection in *G. hirsutum*.

Relatively more QTLs for fiber fineness (Micronaire) than other traits were identified in the present study, similar to an Upland cotton intraspecific population (Shappley et al. 1998) and a *G. hirsutum* × *G. barbadense* BC<sub>1</sub>/BC<sub>2</sub> population (Lacape et al. 2005), but different from another *G. barbadense* introgression population which showed more QTLs for fiber length (Chee et al. 2005b). Rong et al. (2007) summarized fiber quality QTLs from ten (*G. hirsutum* × *G. barbadense*) populations and found more QTLs for fiber fineness than any other trait, although fiber elongation and fiber length showed more QTLs than the remaining traits.

#### Common QTLs shared by different populations

Among the 28 QTLs detected in the present study, eight (28.6%) controlling fiber quality traits were also found in the same chromosome regions (either associated with a same RFLP marker or overlapped by common marker) in previous studies with different populations (Paterson et al. 2003; Draye et al. 2005). These common QTLs included seven (qFF05.1, qFF13.1, qFF14.1, qFL08.1, qFL08.2, qFS22.1 and qFU14.1) identified in *G. hirsutum* × *G. barbadense* F<sub>2</sub> populations (Paterson et al. 2003) and one (qFF05.1) in a *G. barbadense* introgression BC<sub>3</sub>F<sub>2</sub> population (Draye et al. 2005). These common QTLs explained less than 10% of the phenotypic variance in the present study, reiterating prior findings that the advanced-backcross population may be particularly effective for mapping small-effect QTLs that escape detection in other studies (Chee et al. 2005a).

#### Relationships between QTLs for fiber quality traits

Previous studies showed that many QTLs controlling fiber quality traits often co-located on some chromosomal regions in allotetraploid cotton (Saranga et al. 2002; Paterson et al. 2003; Chee et al. 2005b; Lacape et al. 2005; Rong et al. 2007; Wan et al. 2007; Zhang et al. 2009). In the present *G. tomentosum* introgression populations, QTLs for fiber elongation corresponded with three QTLs for fiber fineness (qFE14.1 = qFF14.1, qFE19.2 = qFF19.1, qFE21.2 = qFF21.1) and one for fiber length uniformity (qFE14.1 = qFU14.1). QTLs for fiber fineness corresponded with one QTL for fiber length (qFF04.1 = qFL04.1) and one for fiber length uniformity (qFF14.1 = qFU14.1). QTLs for fiber

strength corresponded with two QTLs for fiber length uniformity (qFS15.1 = qFU15.1, qFS22.1 = qFU22.1). These co-locating QTLs were partially responsible for the correlations among traits. For example, the two favorable alleles of fiber strength and fiber length uniformity present on the same two chromosomal regions contributed to the significant positive relationship between fiber strength and fiber length uniformity. Similar relationships were also found among fiber length-related traits in a *G. barbadense* introgression population (Chee et al. 2005b).

#### QTL interaction between genotype and background

In the present study, one *G. tomentosum* introgression region on Chr. 5, which segregated among more than two family lines in CA3084 background population, showed a significant genotype × family interaction for fiber strength, but the QTL main effect was not significant. Similar interactions between QTL and genetic backgrounds were observed in Upland cotton population with *G. barbadense* chromosome segments (Chee et al. 2005a, b; Draye et al. 2005) and in other crops (Bernacchi et al. 1998; Lecomte et al. 2004). A possible explanation for this observation is that some chromosomal regions may harbor linked alleles with opposite effects (Bernacchi et al. 1998). When the linkages were broken in some of the BC<sub>3</sub>F<sub>2</sub> families, the QTLs may show either positive or negative effects depending on which allele is present. Alternatively, the effects of these QTLs may depend on the presence of genetic loci from other donor chromosome segments (Chee et al. 2005a). For the 28 common markers segregating in both CA3084 and CA3093 background population, two significant genotype × background interactions were identified. This result showed that the different *G. hirsutum* cultivars carried different alleles for fiber quality and yield traits, or different alleles at trans-acting loci influencing these QTLs.

#### QTL interaction between genotype and environment

The extent to which the performance of complex traits varies among environments reflects the complexity of genotype × environment interactions (Saranga et al. 2002; Paterson et al. 2003). In the present study, significant environmental variances were found for all fiber quality traits, and many significant environment effects were identified in the two populations. Only one fiber strength locus, on Chr. 15 in the CA3084 background population, showed significant genotype × environment interaction. This result showed that although many QTLs controlling fiber quality traits had significant environment effects, only a few QTLs had significant genotype × environment interactions.

## QTL distribution across subgenomes

This study adds further evidence to prior studies, indicating that the majority of genetic variation associated with fiber quality in tetraploid cotton traces to the D-subgenome from a diploid ancestor that does not produce spinnable fiber (Jiang et al. 1998; Rong et al. 2007). Of the 28 QTLs, 18 (64.3%) claimed were on D-subgenome chromosomes, albeit with individual traits varying in their distribution from 100% (of six, fiber uniformity) D-subgenome to 100% (of one, weight per boll) A-subgenome. Favorable alleles from *G. tomentosum* were found for six of ten A-subgenome and 12 of 18 D-subgenome loci, suggesting that the evolutionary forces resulting in these favorable alleles apply similarly to the two subgenomes.

## Synthesis

The present study showed that *G. tomentosum* contains many favorable alleles for fiber quality traits, offering a potentially important means to improve Upland cotton. Especially attractive are some favorable alleles that are linked together and might be co-introgressed. Although barriers to gene introgression from other allotetraploid species may exist (Stephens 1949; Jiang et al. 2000), the availability of DNA markers linked to QTLs identified in this and other studies (Paterson et al. 2003; Chee et al. 2005a, b; Draye et al. 2005; Lacape et al. 2005) may assist breeders in transferring and maintaining these traits during Upland cultivar development. Furthermore, because many QTLs from interspecific introgression are in a near-isogenic state in the advanced backcross population, the phenotypic effect measured for each QTL is likely to be a better predictor of its ultimate effect when transferred to other cultivated backgrounds (Chee et al. 2005b). The present QTL mapping offers an additional source of allelic variation to Upland cotton germplasm.

The levels and patterns of segregation distortion in these populations, briefly noted herein and further investigated in a companion study (Waghmare et al., submitted), have consequences for how *G. tomentosum* introgressions might best be deployed in crop improvement. As noted above, among those codominant loci that remain heterozygous in BC<sub>3</sub>F<sub>1</sub>, in BC<sub>3</sub>F<sub>2</sub> heterozygotes occurred at an average frequency of 28.9% whereas *G. tomentosum* homozygotes occurred at 1.8%. This suggests that *G. tomentosum* introgressions might be especially attractive for use in F<sub>1</sub> hybrid cottons, if genotypes homozygous for the introgressions can be stabilized. The use of F<sub>1</sub> hybrids often mitigates ‘linkage drag’ associated with introgression, by virtue of the presence of adapted alleles from one parent. Natural selection against *G. tomentosum* introgressions, once they are allowed to segregate,

provides a strong incentive to use F<sub>1</sub> seed each year in that the benefits of the introgressions will quickly be lost in F<sub>2</sub> seed from the F<sub>1</sub> plants.

**Acknowledgments** We thank the members of the Paterson laboratory for many valuable contributions, and the US Department of Agriculture (02-01412), US National Science Foundation (IIP-0917856), and Cotton, Inc. for financial support.

## References

- Abdalla AM, Reddy OUK, El-Zik KM, Pepper AE (2001) Genetic diversity and relationships of diploid and tetraploid cottons revealed using AFLP. *Theor Appl Genet* 102:222–229
- Alford BB, Liepa GU, Vanbeber AD (1996) Cottonseed protein: what does the future hold? *Plant Food Hum Nutr* 49:1–11
- Beasley JO (1940) The production of polyploids in *Gossypium*. *J Hered* 31:39–48
- Beasley JO (1942) Meiotic chromosome behaviour in species, species hybrids, haploids and induced polyploids of *Gossypium*. *Genetics* 27:25–54
- Bernacchi D, Beck-Bunn T, Emmatty D, Eshed Y, Inai S, Lopez J, Petiard V, Sayama H, Uhlig J, Zamir D, Tanksley S (1998) Advanced backcross QTL analysis of tomato. II. Evaluation of near-isogenic lines carrying single-donor introgressions for desirable wild QTL-alleles derived from *Lycopersicon hirsutum* and *L. pimpinellifolium*. *Theor Appl Genet* 97:170–180
- Brubaker CL, Wendel JF (1994) Reevaluating the origin of domesticated cotton (*Gossypium hirsutum*; Malvaceae) using nuclear restriction fragment length polymorphisms (RFLPs). *Am J Bot* 81:1309–1326
- Brubaker CL, Koontz JA, Wendel JF (1993) Bidirectional cytoplasmic and nuclear introgression in the New World cottons, *Gossypium barbadense* and *G. hirsutum* (Malvaceae). *Am J Bot* 80:1203–1208
- Cantrell RD, Davis DD (1993) Characterization of *G. hirsutum* × *G. barbadense* breeding lines using molecular markers. *Proc Beltwide Production Conf.* pp 1551–1553
- Chee P, Draye X, Jiang C, Decanini L, Delmonte T, Bredhauer B, Smith CW, Paterson AH (2005a) Molecular dissection of interspecific variation between *Gossypium hirsutum* and *G. barbadense* (cotton) by a backcross-self approach: I. Fiber elongation. *Theor Appl Genet* 111:757–763
- Chee P, Draye X, Jiang CX, Decanini L, Delmonte T, Bredhauer R, Smith CW, Paterson AH (2005b) Molecular dissection of interspecific variation between *Gossypium hirsutum* and *Gossypium barbadense* (cotton) by a backcross-self approach: III. Fiber length. *Theor Appl Genet* 111:772–781
- Chen ZJ, Scheffler BE, Dennis E, Triplett BA, Zhang TZ, Guo WZ, Chen XY, Stelly DM, Rabinowicz PD, Town CD, Arioli T, Brubaker C, Cantrell RG, Lacape JM, Ulloa M, Chee P, Gingle AR, Haigler CH, Percy R, Saha S, Wilkins T, Wright RJ, Deynze AV, Zhu YX, Yu SX, Abdurakhmonov I, Katageri I, Kumar PA, Rahman M, Zafar Y, Yu JZ, Kohel RJ, Wendel JF, Paterson AH (2007) Toward sequencing cotton (*Gossypium*) genomes. *Plant Physiol* 145:1303–1310
- Cherry JP, Leffler HR (1984) Seed. In: Kohel RJ, Lewis CF (eds) *Cotton*. ASA, Madison, Wisconsin, pp 511–569
- Culp TW, Harrell DC, Kerr T (1979) Some genetic implications in the transfer of high fiber strength genes to Upland cotton. *Crop Sci* 19:481–484
- Draye X, Chee P, Jiang CX, Decanini L, Delmonte T, Bredhauer R, Smith CW, Paterson AH (2005) Molecular dissection of interspecific variation between *Gossypium hirsutum* and *Gossypium barba-*



- dense* (cotton) by a backcross-self approach: II. Fiber fineness. *Theor Appl Genet* 111:764–771
- Fryxell PA (1979) The natural history of the cotton tribe. Texas A and M University Press, College Station, TX
- Fryxell PA, Craven LA, Stewart J McD (1992) A revision of *Gossypium* sect. *Grandicalyx* (Malvaceae), including the description of six new species. *Syst Bot* 17:91–114
- Guo WZ, Zhang TZ, Pan JJ, Wang XY (1997) A preliminary study on genetic diversity of Upland cotton cultivars in China. *Acta Gossypii Sinica* 9:242–247
- Iqbal MJ, Aziz N, Saeed NA, Zafar Y, Malik KA (1997) Genetic diversity evaluation of some elite cotton varieties by RAPD analysis. *Theor Appl Genet* 95:139–144
- Iqbal MJ, Reddy OUK, El-Zik KM, Pepper AE (2001) A genetic bottleneck in the 'evolution under domestication' of Upland cotton *Gossypium hirsutum* L. examined using DNA fingerprinting. *Theor Appl Genet* 103:547–554
- Jiang C, Wright R, El-Zik K, Paterson AH (1998) Polyploid formation created unique avenues for response to selection in *Gossypium* (cotton). *Proc Nat Acad Sci USA* 95:4419–4424
- Jiang C, Chee P, Draye X, Morrell P, Smith C, Paterson A (2000) Multi-locus interactions restrict gene flow in advanced-generation interspecific populations of polyploid *Gossypium* (Cotton). *Evolution* 54:798–814
- Khan SA, Hussain D, Askari E, Stewart JM, Malik KA, Zafar Y (2000) Molecular phylogeny of *Gossypium* species by DNA fingerprinting. *Theor Appl Genet* 101:931–938
- Lacape JM, Nguyen TB, Courtois B, Belot JL, Giband M, Gourlot JP, Gawryziak G, Roques S, Hau B (2005) QTL analysis of cotton fiber quality using multiple *Gossypium hirsutum* × *Gossypium barbadense* backcross generations. *Crop Sci* 45:123–140
- Lacape JM, Dessauw D, Rajab M, Noyer JL, Hau B (2007) Microsatellite diversity in tetraploid *Gossypium* germplasm: assembling a highly informative genotyping set of cotton SSRs. *Mol Breeding* 19:45–58
- Lecomte L, Duffe P, Buret M, Servin B, Hospital F, Causse M (2004) Marker-assisted introgression of five QTLs controlling fruit quality traits into three tomato lines revealed interactions between QTLs and genetic backgrounds. *Theor Appl Genet* 109:658–668
- Linos AA, Bebeli PJ, Kaltsikes PJ (2002) Cultivar identification in Upland cotton using RAPD markers. *Aust J Agric Res* 53:637–642
- Lu HJ, Myers GO (2002) Genetic relationships and discrimination of ten influential Upland cotton varieties using RAPD markers. *Theor Appl Genet* 105:325–331
- Lusas EW, Jividen GM (1987) Glandless cottonseed: a review of the first 25 years of processing and utilization research. *J Am Oil Chem Soc* 64:839–854
- McCouch SR, Cho YG, Yano PE, Blinstrub M, Morishima H, Kinoshita T (1997) Report on QTL nomenclature. *Rice Genet Newslett* 14:11–13
- Multani DS, Lyon BR (1995) Genetic fingerprinting of Australian cotton cultivars with RAPD markers. *Genome* 38:1005–1008
- Paterson AH, DeVerna J, Lanini B, Tanksley SD (1990) Fine mapping of quantitative trait loci using selected overlapping recombinant chromosomes, from an interspecies cross of tomato. *Genetics* 124:735–742
- Paterson AH, Saranga Y, Menz M, Jiang C, Wright RJ (2003) QTL analysis of genotype × environmental interactions affecting cotton fiber quality. *Theor Appl Genet* 106:384–396
- Percival AE, Kohel RJ (1990) Distribution, collection, and evaluation of *Gossypium*. *Adv Agronomy* 44:225–256
- Percival AE, Wendel JF, Stewart JM (1999) Taxonomy and germplasm resources. In: Smith CW, Cothren JT (eds) Cotton origin, history, technology, and production. John Wiley and Sons Inc., New York, pp 33–63
- Percy RG, Wendel JF (1990) Allozyme evidence for the origin and diversification of *Gossypium barbadense* L. *Theor Appl Genet* 79:529–542
- Rong J, Feltus FA, Waghmare VN, Pierce GJ, Chee PW, Draye X, Saranga Y, Wright RJ, Wilkins TA, May OL, Smith CW, Gannaway JR, Wendel JF, Paterson AH (2007) Meta-analysis of polyploid cotton QTL shows unequal contributions of subgenomes to a complex network of genes and gene clusters implicated in lint fiber development. *Genetics* 176:2577–2588
- Rungis D, Llewellyn D, Dennis ES, Lyon BR (2005) Simple sequence repeat (SSR) markers reveal low levels of polymorphism between cotton (*Gossypium hirsutum* L.) cultivars. *Aus J Agr Res* 56:301–307
- Saha S, Wu J, Jenkins JN, Mccarty JC, Gutierrez OA, Stelly DM, Percy RG, Raska DA (2004) Effect of chromosome substitutions from *Gossypium barbadense* L. 3–79 into *G. hirsutum* L. TM-1 on agronomic and fiber traits. *J Cotton Sci* 8:162–169
- Saha S, Jenkins JN, Wu J, Mccarty JC, Gutierrez OA, Percy RG, Cantrell RG, Stelly DM (2006) Effects of chromosome-specific introgression in Upland cotton on fiber and agronomic traits. *Genetics* 172:1927–1938
- Saranga Y, Menz M, Jiang CX, Wright RJ, Yakir D, Paterson AH (2002) Genetic mapping implicates osmotic potential as a major component of crop adaptation to arid conditions. *Genome Res* 11:1988–1995
- SAS Institute Inc. (2008) SAS/STAT<sup>®</sup> 9.2 User's Guide. SAS Institute Inc, Cary
- Self SG, Liang KL (1987) Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *J Am Stat Assoc* 82:605–610
- Shappley ZW, Jenkins JN, Zhu J, Jack CM (1998) Quantitative trait loci associated with agronomic and fiber traits of Upland cotton. *J Cotton Sci* 2:153–163
- Stephens SG (1949) The cytogenetics of speciation in *Gossypium*. I. Selective elimination of the donor parent genotype in interspecific backcrosses. *Genetics* 34:627–637
- Tanksley SD, Nelson CJ (1996) Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. *Theor Appl Genet* 92:191–203
- Van Ooijen JW (2009) MapQTL 6.0. Software for the mapping of quantitative trait loci in experimental populations. Plant Research International, Wageningen
- Voorrips RE (2006) MapChart 2.2: software for the graphical presentation of linkage maps and QTLs. Plant Research International, Wageningen
- Vroh Bi I, Baudoin JP, Mergeai G (1998) Cytogenetics of the glandless-seed and glanded-plant trait from *Gossypium sturtianum* Willis introgressed into Upland cotton (*Gossypium hirsutum* L.). *Plant Breeding* 117:235–241
- Vroh Bi I, Baudoin JP, Hau B, Mergeai G (1999a) Development of high-gossypol cotton plants with low-gossypol seeds using trispecies bridge crosses and in vitro culture of seed embryos. *Euphytica* 106:243–251
- Vroh Bi I, Maquet A, Baudoin JP, du Jardin P, Jacquemin JM, Mergeai G (1999b) Breeding for 'low-gossypol seed and high-gossypol plants' in Upland cotton. Analysis of trispecies hybrids and backcross progenies using AFLPs and mapped RFLPs. *Theor Appl Genet* 99:124–1233
- Waghmare VN, Rong JK, Rogers CJ, Pierce GJ, Wendel JF, Paterson AH (2005) Genetic mapping of a cross between *Gossypium hirsutum* (cotton) and the Hawaiian endemic, *Gossypium tomentosum*. *Theor Appl Genet* 111:665–676
- Wan Q, Zhang ZS, Hu MC, Chen L, Liu DJ, Chen X, Wang W, Zheng J (2007) T<sub>1</sub> locus in cotton is the candidate gene affecting lint per-

- centage. fiber quality and spiny bollworm (*Earias* spp.) resistance. *Euphytica* 158:241–247
- Wendel JF, Brubaker CL, Percival AE (1992) Genetic diversity in *Gossypium hirsutum* and the origin of Upland cotton. *Am J Bot* 79:1291–1310
- Xiao J, Grandillo S, Ahn S, Yuan L, Tanksley SD, McCouch SR (1996) Genes from wild rice improve yield. *Nature* 384:223–224
- Xu QH, Zhang XL, Nie YC (2001) Genetic diversity evaluation of cultivars (*G. hirsutum* L.) from the Changjiang River valley and Yellow River valley by RAPD markers. *Acta Genet Sin* 28:683–690 (in Chinese with English abstract)
- Young ND, Tanksley SD (1989) RFLP analysis of the size of chromosomal segments retained around the Tm-2 locus of tomato during backcross breeding. *Theor Appl Genet* 77:353–359
- Zhang ZS, Hu MC, Zhang J, Liu DJ, Zheng J, Zhang K, Wang W, Wan Q (2009) Construction of a comprehensive PCR-based marker linkage map and QTL mapping for fiber quality traits in (*Gossypium hirsutum* L.). *Mol Breeding* 24:49–61
- Zhu LF, Zhang XL, Nie YC (2003) Analysis of genetic diversity in Upland cotton (*Gossypium hirsutum* L.) cultivars from China and foreign countries by RAPDs and SSRs. *J Agric Biotechnol* 11:450–455 (in Chinese with English abstract)