Theor Appl Genet (2017) 130:1297–1308
DOI 10.1007/s00122-017-2889-1

ORIGINAL ARTICLE

QTL analysis of cotton fiber length in advanced backcross populations derived from a cross between *Gossypium hirsutum* and *G. mustelinum*

Baohua Wang[1,2,3] · Xavier Draye[4] · Zhimin Zhuang[2,3] · Zhengsheng Zhang[1,5] · Min Liu[1] · Edward L. Lubbers[3] · Don Jones[6] · O. Lloyd May[3,7] · Andrew H. Paterson[1] · Peng W. Chee[3]

Received: 30 June 2016 / Accepted: 2 March 2017 / Published online: 27 March 2017

Abstract

Key message **QTLs for fiber length mapped in three generations of advanced backcross populations derived from crossing *Gossypium hirsutum* and *Gossypium mustelinum* showed opportunities to improve elite cottons by introgression from wild relatives.**

Abstract The molecular basis of cotton fiber length in crosses between *Gossypium hirsutum* and *Gossypium mustelinum* was dissected using 21 BC_3F_2 and 12 corresponding $BC_3F_{2:3}$ and $BC_3F_{2:4}$ families. Sixty-five quantitative trait loci (QTLs) were detected by one-way analysis of variance. The QTL numbers detected for upper-half mean length (UHM), fiber uniformity index (UI), and short fiber content (SFC) were 19, 20, and 26 respectively. Twenty-three of the 65 QTLs could be detected at least twice near adjacent markers in the same family or near the same markers across different families/generations, and 32 QTLs were detected in both one-way variance analyses and mixed model-based composite interval mapping. *G. mustelinum* alleles increased UHM and UI and decreased SFC for five, one, and one QTLs, respectively. In addition to the main-effect QTLs, 17 epistatic QTLs were detected which helped to elucidate the genetic basis of cotton fiber length. Significant among-family genotypic effects were detected at 18, 16, and 16 loci for UHM, UI, and SFC, respectively. Six, two, and two loci showed genotype × family interaction for UHM, UI and SFC, respectively, illustrating complexities that might be faced in introgression of exotic germplasm into cultivated cotton. Co-location of many QTLs for UHM, UI, and SFC accounted for correlations among these traits, and selection of these QTLs may improve the three traits simultaneously. The simple sequence repeat (SSR) markers associated with *G. mustelinum* QTLs will assist breeders in transferring and maintaining valuable traits from this exotic source during cultivar development.

Communicated by Joshua A. Udall.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-017-2889-1) contains supplementary material, which is available to authorized users.

✉ Andrew H. Paterson
paterson@uga.edu

✉ Peng W. Chee
pwchee@uga.edu

1 Plant Genome Mapping Laboratory, University of Georgia, 111 Riverbend Road, Athens, GA 30602, USA

2 School of Life Sciences, Nantong University, Nantong 226019, Jiangsu, China

3 Department of Crop and Soil Sciences, University of Georgia, 2356 Rainwater Road, Tifton, GA 31793, USA

4 Université catholique de Louvain, Place Croix du Sud 2/11, 1348 Louvain-la-Neuve, Belgium

5 Engineering Research Center of South Upland Agriculture, Ministry of Education, Southwest University, Chongqing 400716, China

6 Cotton Incorporated, Cary, NC 27513, USA

7 Present Address: Monsanto Cotton Breeding, Tifton, GA 31793, USA

Introduction

This is the third study describing the results of a backcross-self-approach to map and introgress quantitative trait loci (QTLs) for fiber quality traits from *Gossypium mustelinum* into Upland cotton (*G. hirsutum*). In the previous two studies, we described the interspecific genetic linkage map and 24 QTLs for fiber elongation (15 detected in BC_3F_2 and 14

in $BC_3F_{2:3}$ and $BC_3F_{2:4}$ but five are common; Wang et al. 2016a, b). For one-third of these QTLs, the *G. mustelinum* alleles improved fiber elongation, indicating the potential benefit of introgressing genes from this species in breeding for higher quality Upland cotton.

In this study, we report the results of interspecific introgression and QTL analysis for fiber length, which is considered by some to be the most important property of cotton in raw material marketing and yarn processing (Kuang and Yu 2015). The most widely used parameters for fiber length include the average length of the longer half of the fiber span length distribution (i.e., upper-half mean length, UHM); length uniformity or uniformity ratio or fiber uniformity index (UI) which is defined as "a ratio between the mean length and the upper-half mean length of the fibers and is expressed as a percentage" (Testore and Minero 1988); and short fiber content (SFC), which is currently specified as the percentage of fibers shorter than 10, 11, 12, or 12.5 mm in the great majority of cases. Rotor yarn strength initially increases with staple length (fiber length) up to 40 mm and then remains constant, although with coarse fibers there is a slight increase up to 60 mm (Lawrence 2003). Longer fibers can be processed at greater efficiencies and produce finer yarns with higher quality than short fibers, since short fibers require increased twisting during spinning, reducing production speed, and causing changes to other yarn properties that impair fabric properties (Lawrence 2003). Low uniformity index is often accompanied by a high content of short fibers, which reduces the quality of the textile product.

The objectives of the investigation reported in this study were to determine the number, chromosomal locations, and phenotypic effects of QTLs for fiber length in an interspecific *G. hirsutum*×*G. mustelinum* backcross-self-population and to investigate the relationship between QTLs associated with UHM, UI, and SFC. The QTLs identified in this study add a new level of information to our knowledge of the genetic basis of fiber quality, also laying the foundation for marker-assisted introgression of *G. mustelinum* alleles to improve fiber quality of Upland cotton.

Materials and methods

Population development and field evaluation

A modified backcross-self-mating design was used in this study. Advanced-generation backcross populations were developed by first crossing *G. hirsutum* acc. PD94042 and *G. mustelinum* (AD4-8), then independently backcrossing F_1 plants to the *G. hirsutum* parent for three cycles. A total of 21 lineages led to the production of at least one BC_3F_1 plant, each of which was self-pollinated to generate 21 BC_3F_2 families ranging in size from 127 to 160 plants per family (totally 3203 BC_3F_2 progeny) that were planted in 2006 in Tifton Georgia. In addition, 12 $BC_3F_{2:3}$ and $BC_3F_{2:4}$ families of 130–160 lines per family (totally 1826 lines) were derived from 12 of the 21 BC_3F_2 families with enough seeds, and were planted in completely randomized designs with two replicate plots for each line in 2008 and 2009 in Tifton, Georgia. All cultural practices followed standard recommendations for Georgia cotton production as described (Wang et al. 2016a). Seed cotton was hand-harvested from all bolls of each plant for BC_3F_2 family, and from 50 bolls from the middle of each plot (to avoid edge effects) for $BC_3F_{2:3}$ and $BC_3F_{2:4}$ families, and ginned on a saw gin. For phenotyping, the fiber quality data of each individual were collected for BC_3F_2, whereas the fiber quality data of two replicated lines were collected for each line in $BC_3F_{2:3}$ and $BC_3F_{2:4}$ generation. The family mean values of 21 BC_3F_2 families and 12 $BC_3F_{2:3}$ and $BC_3F_{2:4}$ families were used to demonstrate the phenotypic distribution of fiber length traits measured with family mean deviation from the recurrent *Gossypium hirsutum* parent. Three traits reflecting fiber length parameters, specifically UHM, UI, and SFC were determined by the Cotton Incorporated Textile Services Laboratory (Cotton Incorporated, Cary, N.C.) using the High-Volume Precision Instrument (HVI; Zellweger-Uster, Knoxville, Tenn.).

Genotyping and data analysis

A total of 218 SSR markers that were approximately evenly distributed on our *G. hirsutum* by *G. mustelinum* map constructed from an F_2 population of the same two parents comprising 1055 loci (Wang et al. 2016b) were used for genotyping. The markers that detected introgression of *G. mustelinum* alleles in the BC_3F_1 were used to screen every individual in each BC_3F_2 family, which constitutes the genotype of each BC_3F_2 individual and also the corresponding $BC_3F_{2:3}$ and $BC_3F_{2:4}$ lines.

Since many populations segregated for *G. mustelinum* alleles in only small segments of the genome, we tested associations between fiber length phenotypes and marker genotypes for statistical significance by one-way variance analyses for every marker locus segregating within each BC_3 family. The GLM procedure of the SAS ver.8 software package (SAS Institute 1999) was used to perform the analyses, with a significance threshold set at $P<0.001$ (F-test). The appropriate linear combination of the estimated model parameters was used to evaluate the modes of gene action (additive, a; and dominant, d) for individual QTLs, and their significance levels were assessed with corresponding contrasts as described (Paterson et al. 1990). QTLs were considered to be over- or under-dominant if the absolute value of the dominance/additive (d/a) ratio exceeded 3

(Chee et al. 2005a). The R^2 of variance analysis at a marker locus estimated the proportion of phenotypic variance accounted for by a nearby QTL.

In order to confirm the reliability of the QTLs detected by one-way variance analyses and also to map epistatic QTLs, QTLNetwork V2.1 (Yang et al. 2008) was also used to analyze QTLs for fiber length traits in each of the 21 BC_3F_2 families and 12 $BC_3F_{2:3}$ and $BC_3F_{2:4}$ families. The mixed model-based composite interval mapping (MCIM) method was used with the critical F value calculated based on 1000 permutation tests, with window size and walk speed set at 10 and 1 cM, respectively. The significance threshold was set at $P=0.001$ to claim a putative main-effect or epistatic QTL. QTLs for fiber length traits were mapped in each of the 21 BC_3F_2 families and 12 $BC_3F_{2:3}$ and $BC_3F_{2:4}$ families. QTLs were also detected in joint analysis of the 12 families grown in three generations/environments (BC_3F_2, $BC_3F_{2:3}$, and $BC_3F_{2:4}$), also considering environmental effects. QTLs detected by QTLNetwork 2.1 sharing a common marker with those detected by one-way variance analyses were considered to be the same QTLs. QTL nomenclature was as described (McCouch et al. 1997), with the QTL name beginning with "q" indicating a QTL, followed by an abbreviation of the trait name, the name of chromosome (Wang et al. 2016b; Zhang et al. 2015b), and consecutive numbers when more than one QTL was detected on the same chromosome for the same trait. The sequences and genome positions of all the markers that linked to QTLs are available from CottonGen (Yu et al. 2014a; Zhang et al. 2015b).

For loci segregating in two or more families, the MIXED procedure of the SAS ver. 8 package was used to perform two-way mixed model variance analyses, including genotype (G) as a fixed factor and family (F) and genotype × family (G × F) interaction as random factors. The residual maximum likelihood (REML) method was used to estimate model parameters, and an F statistic using a general Satterthwaite approximation for the denominator degrees of freedom (SAS Institute 1999) was applied to test marker-trait association (genotype factor). A likelihood-ratio (ChiSq) test was performed for the G × F interaction (Chee et al. 2005a; Self and Liang 1987). A significance level of $P<0.001$ was set for G and G × F effects.

Results

Phenotypic distribution and correlations

The distributions of fiber quality traits of the BC_3 progenies are shown in Table 1, Supplementary Fig. 1. All traits expressed transgressive segregation in the three generations, with distribution ranges being wider in BC_3F_2 than $BC_3F_{2:3}$ and $BC_3F_{2:4}$ (Table 1). Although *G. mustelinum* does not produce spinnable fiber, many BC_3 progenies have higher UHM and UI, and lower SFC (usually lower SFC is preferred) than the *G. hirsutum* parent (Table 1). For UHM, six of 21 BC_3F_2, four of 12 $BC_3F_{2:3}$, and three of 12 $BC_3F_{2:4}$ families had mean values exceeding that of the *G. hirsutum* parent; three families, namely POP10, POP11, and POP27 showed higher mean values than the *G. hirsutum* parent in all generations. For UI, 15 of 21 BC_3F_2, no $BC_3F_{2:3}$, and only 1 of 12 $BC_3F_{2:4}$ families had mean values exceeding that of the *G. hirsutum* parent. For SFC, 16 of 21 BC_3F_2, but only 2 of 12 families in both $BC_3F_{2:3}$ and $BC_3F_{2:4}$ had lower mean values than the *G. hirsutum* parent (Supplementary Fig. 1). Correlations were calculated to detect relationships between the three fiber length traits. There was significant positive correlation between UHM and UI ($P<0.01$), and significant negative correlation of SFC with both UHM and UI ($P<0.01$) in all three generations (Table 2).

Main effect QTLs detected for each trait

A total of 65 non-overlapping QTLs were estimated to be segregating in the three generations (Table 3). Blocks of linked markers within a family that showed significant

Table 1 Summary statistics of fiber quality traits measured on the *G. hirsutum* parent and BC_3 progenies

Generation	Trait	Progeny			*G. hirsutum*
		Range	Mean ± SD	CV (%)	Parent
BC_3F_2	Upper-half mean length (UHM, mm)	20.1–35.1	28.5 ± 2.26	7.9	28.8
	Fiber uniformity index (UI)	74.8–87.9	83.8 ± 1.73	2.1	83.2
	Short fiber content (SFC)	5.8–15.4	8.0 ± 1.26	15.8	8.3
$BC_3F_{2:3}$	Upper-half mean length (UHM, mm)	21.8–33.5	28.9 ± 1.74	6.0	29.7
	Fiber uniformity index (UI)	76.9–86.3	83.2 ± 1.25	1.5	84.5
	Short fiber content (SFC)	6.1–13.4	7.6 ± 0.79	10.4	7.3
$BC_3F_{2:4}$	Upper-half mean length (UHM, mm)	25.4–33.3	29.6 ± 1.27	4.3	29.7
	Fiber uniformity index (UI)	80.5–86.2	83.6 ± 0.85	1.0	84.4
	Short fiber content (SFC)	6.2–9.4	7.3 ± 0.43	5.9	7.1

Table 2 Correlation coefficients between different fiber quality traits based on mean values of three generations

Generation	Trait	UHM	UI
BC_3F_2	UI	0.427**	
	SFC	−0.319**	−0.861**
$BC_3F_{2:3}$	UI	0.595**	
	SFC	−0.306**	−0.768**
$BC_3F_{2:4}$	UI	0.583**	
	SFC	−0.450**	−0.703**

**Significant at the 0.01 level

marker-trait association ($P < 0.001$) were inferred to represent single QTLs. These QTLs were mapped to 19 chromosomes with 26 on eight A-subgenome chromosomes, and 39 on 11 D-subgenome chromosomes. In 19 of the 21 families we detected one or more QTLs for fiber length traits, with a maximum of eight QTLs in family POP02. Twenty-three of the 65 QTLs could be detected at least twice near linked markers in the same family or near the same markers in different families/generations. For example, as shown in Table 3, the QTL *qUHM-5-1* was detected near two markers (BNL2656 and NAU3498) in the same family, POP15; another QTL (*qUHM-24-1*) was detected near the same marker, DPL0068, in two different families, POP16 and POP20.

Twelve QTLs explained more than 20% of PV, and four could be detected at least twice near linked markers in the same family or near the same markers in different families/generations. Thirty-two of the 65 QTLs were also detected by the MCIM method of QTLNetwork. The biometrical parameters of all significant marker-trait associations are listed in Table 3. A summary of the QTLs detected for each trait follows.

QTLs for UHM

A total of 19 non-overlapping QTLs were detected on 14 chromosomes for UHM (Table 3), with five in the A-subgenome and 14 in the D-subgenome. Eleven were found at least twice near linked markers in the same family or near the same markers in different families/generations, namely *qUHM-3-1, qUHM-5-1, qUHM-7-1, qUHM-8-1, qUHM-14-1, qUHM-19-2, qUHM-21-1, qUHM-24-1, qUHM-24-2, qUHM-25-1*, and *qUHM-26-1*. The PVE per individual association ranged from 9.05% (*qUHM-19-2*) to 32.01% (*qUHM-24-2*), with an average of 16.06%; and five QTLs explained more than 20% of PV (*qUHM-8-1, qUHM-11-1, qUHM-23-1, qUHM-24-2, qUHM-26-1*). For 14 of the 19 QTLs, alleles from *G. hirsutum* increased UHM, which was consistent with the parental phenotypes. For five of the 19 QTLs, alleles from *G. mustelinum* increased UHM (Table 3). Twelve of the 19 QTLs were also detected by the MCIM method of QTLNetwork (Table 3).

QTLs for UI

A total of 20 non-overlapping QTLs for UI were mapped to 14 chromosomes with ten mapped to seven A-subgenome chromosomes and ten to seven D-subgenome chromosomes. Six QTLs were found at least twice near linked markers in the same family or near the same markers in different families/generations. The PVE per individual locus ranged from 8.52% (*qUI-5-1*) to 26.48% (*qUI-24-1*), with an average of 14.66%; and five QTLs explained more than 20% of PV (*qUI-11-1, qUI-12-2, qUI-12-3, qUI-24-1, qUI-26-1*). For 19 of the 20 QTLs, *G. hirsutum* alleles increased UI, consistent with the parental phenotypes. For one of the 20 QTLs (*qUI-14-1*), the *G. mustelinum* allele increased UI (Table 3). Ten of the 20 QTLs were also detected by the MCIM method of QTLNetwork (Table 3). Among these 10 stringent QTLs, four explained more than 20% of PV; one (*qUI-24-1*) was found in two generations in the same family.

QTLs for SFC

A total of 26 non-overlapping QTLs of SFC were mapped to 13 chromosomes; eleven to six A-subgenome chromosomes and 15 to seven D-subgenome chromosomes. Six QTLs were detected at least twice near linked markers in the same family or near the same markers in different families/generations. The PVE per individual locus ranged from 8.56% (*qSFC-7-2*) to 21.26% (*qSFC-24-3*), with an average of 13.77%; and two QTLs explained more than 20% of PV. For 25 QTLs, *G. mustelinum* alleles increased SFC, consistent with the parental phenotypes. For one of the 26 QTLs (*qSFC-17-2*), the *G. mustelinum* allele decreased SFC (Table 3). Ten of the 26 QTLs were also detected by the MCIM method of QTLNetwork (Table 3).

Co-location of main-effect QTLs for different traits

Sixteen cases of co-location of QTLs for different traits suggest that marker-assisted selection based on these QTLs may improve each of three measures of fiber length simultaneously (Table 3). Co-location of QTLs for UHM and UI were found in four regions, namely *qUHM-5-1* and *qUI-5-1* near BNL2656 and NAU3498; *qUHM-8-1* and *qUI-8-1* near CIR354b and NAU4900; *qUHM-23-3* and *qUI-23-2* near BNL3511; and *qUHM-23-1* and *qUI-23-1* near DPL0378. Two cases of QTL co-location were detected for UHM and SFC, namely *qUHM-17-1* and *qSFC17-1* near BNL2496A; and *qUHM-19-3* and *qSFC-19-2* near DPL0140. Five cases of QTL co-location

Table 3 Biometrical parameters of QTLs affecting fiber length parameters

QTL[a]	Generation	Locus	Family	R^2 (%)[b]	A[b]	D[b]	D/A ratio[b]	Mode of action[c]
*qUHM-3-1**	$BC_3F_{2:3}$	BNL3267a	POP12	9.43	−0.65	0.06	−0.10	A
	$BC_3F_{2:4}$	BNL3267a	POP12	10.00	−0.55	0.09	0.16	A
	$BC_3F_{2:4}$	DPL0605	POP12	11.00	−0.59	0.06	0.10	A
	$BC_3F_{2:4}$	DPL0354	POP12	14.00	−0.55	0.30	0.55	A
*qUHM-5-1**	$BC_3F_{2:3}$	BNL2656	POP15	9.88	0.72	0.49	0.68	A
	$BC_3F_{2:3}$	NAU3498	POP15	9.30	0.73	−0.01	−0.01	A
*qUHM-7-1**	BC_3F_2	MUCS616	POP12	9.57	0.94	0.32	0.34	A
	$BC_3F_{2:3}$	MUCS616	POP12	9.35	0.70	−0.03	−0.04	A
	BC_3F_2	MUCS616	POP16	11.52	1.41	1.22	0.87	A
	$BC_3F_{2:3}$	NAU2002	POP12	9.10	0.95	–	–	–
qUHM-8-1	BC_3F_2	CIR354b	POP31	25.51	1.67	0.74	0.44	A
	BC_3F_2	NAU4900	POP31	23.82	1.46	0.49	0.34	A
qUHM-11-1	BC_3F_2	BNL3442	POP05	23.07	1.36	−0.63	−0.46	A
qUHM-14-1	BC_3F_2	BNL226a	POP10	21.03	−1.50	–	–	–
	$BC_3F_{2:3}$	BNL226a	POP10	10.00	−0.86	–	–	–
qUHM-15-1	BC_3F_2	NAU4045	POP15	9.15	−2.74	–	–	–
*qUHM-17-1**	BC_3F_2	BNL2496A	POP02	11.42	0.77	0.66	0.86	A
qUHM-19-1	BC_3F_2	BNL285	POP10	15.91	0.86	1.77	2.06	D
*qUHM-19-2**	BC_3F_2	NAU5489	POP12	9.05	0.88	0.55	0.62	A
	$BC_3F_{2:3}$	NAU5489	POP12	11.47	0.77	0.27	0.34	A
*qUHM-19-3**	$BC_3F_{2:3}$	DPL0140	POP35	13.81	−0.60	0.42	-0.69	A
*qUHM-21-1**	BC_3F_2	BNL1034	POP16	16.08	1.15	−0.57	−0.50	A
	BC_3F_2	BNL2589	POP31	24.01	1.84	−0.23	−0.12	A
	BC_3F_2	BNL3171	POP16	15.20	0.98	−0.87	−0.89	A
	BC_3F_2	BNL3171	POP31	22.81	1.70	0.37	0.22	A
	$BC_3F_{2:4}$	NAU3074	POP16	9.94	0.41	−0.02	−0.04	A
qUHM-23-1	BC_3F_2	DPL0378	POP05	21.99	1.42	0.01	0.01	A
*qUHM-23-2**	BC_3F_2	BNL3383	POP10	13.36	0.85	0.46	0.54	A
*qUHM-23-3**	$BC_3F_{2:3}$	BNL3511	POP31	19.07	1.34	0.80	0.59	A
*qUHM-24-1**	BC_3F_2	DPL0068	POP16	15.10	1.74	0.37	0.22	A
	BC_3F_2	DPL0068	POP20	14.02	1.04	0.06	0.06	A
*qUHM-24-2**	BC_3F_2	NAU3605	POP31	28.11	1.76	0.73	0.41	A
	$BC_3F_{2:3}$	NAU3605	POP31	19.77	1.34	0.35	0.26	A
	BC_3F_2	TMHB1	POP04	16.23	1.89	1.03	0.54	A
	BC_3F_2	TMHB1	POP31	32.01	1.82	0.24	0.13	A
	$BC_3F_{2:3}$	TMHB1	POP31	16.66	1.20	0.25	0.20	A
*qUHM-25-1**	$BC_3F_{2:3}$	BNL3264	POP17	15.64	−0.25	0.59	−2.34	D
	$BC_3F_{2:4}$	BNL3264	POP17	14.88	−0.58	−0.05	0.09	A
qUHM-26-1	BC_3F_2	NAU3860	POP16	25.01	0.92	3.02	3.27	H
	$BC_3F_{2:3}$	NAU3860	POP16	14.11	0.30	1.23	4.14	H
	BC_3F_2	STV122	POP16	27.25	0.35	2.39	6.77	H
qUI-2-1	BC_3F_2	BNL1897	POP02	10.44	0.54	0.23	0.43	A
	BC_3F_2	BNL3512a	POP02	11.33	0.58	0.36	0.62	A
*qUI-5-1**	$BC_3F_{2:3}$	BNL2656	POP15	9.42	0.51	0.38	0.73	A
	$BC_3F_{2:3}$	NAU3498	POP15	8.52	0.53	0.07	0.12	A
*qUI-5-2**	BC_3F_2	BNL3029	POP04	10.28	0.95	–	–	–
*qUI-7-1**	$BC_3F_{2:3}$	MUCS616	POP12	9.32	0.52	0.06	0.12	A
	$BC_3F_{2:3}$	NAU2002	POP12	10.58	0.75	–	–	–
qUI-8-1	$BC_3F_{2:3}$	CIR354b	POP31	16.94	0.98	0.84	0.86	A

Table 3 (continued)

QTL[a]	Generation	Locus	Family	R^2 (%)[b]	A[b]	D[b]	D/A ratio[b]	Mode of action[c]
	$BC_3F_{2:3}$	NAU4900	POP31	15.94	0.90	0.71	0.79	A
qUI-10-1	$BC_3F_{2:3}$	JESPR6	POP35	10.62	0.18	0.74	4.04	H
qUI-11-1	BC_3F_2	BNL1151	POP31	25.77	1.32	−0.16	−0.12	A
qUI-12-1	BC_3F_2	JESPR300	POP06	11.97	0.53	0.17	0.32	A
*qUI-12-2**	BC_3F_2	CIR293	POP32	24.47	1.60	0.15	0.10	A
*qUI-12-3**	BC_3F_2	DPL0866	POP11	22.27	1.32	1.04	0.79	A
qUI-14-1	BC_3F_2	STS236	POP03	15.78	−0.22	1.17	−5.42	H
qUI-15-1	BC_3F_2	BNL1350	POP02	14.93	0.78	0.67	0.85	A
	BC_3F_2	BNL2646	POP02	10.67	0.70	0.71	1.01	D
	BC_3F_2	BNL2700	POP02	10.93	0.68	0.53	0.77	A
qUI-18-1	BC_3F_2	MUSS603	POP04	9.37	0.84	0.52	0.62	A
qUI-19-1	$BC_3F_{2:3}$	NAU5489	POP12	9.15	0.51	0.19	0.39	A
*qUI-23-1**	BC_3F_2	DPL0378	POP35	15.43	1.06	0.20	0.19	A
qUI-23-2	$BC_3F_{2:3}$	BNL3511	POP31	18.60	0.86	0.75	0.88	A
*qUI-24-1**	BC_3F_2	TMHB1	POP31	26.48	1.16	0.48	0.42	A
	$BC_3F_{2:3}$	TMHB1	POP31	14.50	0.78	0.13	0.16	A
*qUI-24-2**	BC_3F_2	NAU3605	POP02	18.03	0.89	0.22	0.25	A
*qUI-26-1**	BC_3F_2	NAU3862	POP02	21.14	0.72	−0.31	−0.43	A
*qUI-26-2**	BC_3F_2	STV122	POP16	12.99	0.44	0.95	2.15	D
qSFC-2-1	BC_3F_2	BNL3971	POP04	9.33	−0.81	−0.65	0.80	A
*qSFC-2-2**	$BC_3F_{2:3}$	BNL1434	POP17	12.77	−0.35	0.25	−0.70	A
qSFC-3-1	$BC_3F_{2:4}$	DPL0354	POP35	16.62	−0.38	−0.36	0.93	A
qSFC-6-1	BC_3F_2	NAU5433	POP32	17.29	−1.34	−0.96	0.72	A
qSFC-7-1	BC_3F_2	BNL1604	POP01	15.27	−0.36	−0.53	1.48	D
	BC_3F_2	NAU1305	POP01	13.39	−0.37	−0.52	1.42	D
	BC_3F_2	NAU1305	POP27	12.15	−0.59	−0.05	0.09	A
*qSFC-7-2**	$BC_3F_{2:3}$	MUCS616	POP12	11.80	−0.41	−0.03	0.07	A
	$BC_3F_{2:4}$	MUCS616	POP12	11.00	−0.23	−0.06	0.26	A
	$BC_3F_{2:3}$	MUSS006	POP12	8.56	−0.36	−0.12	0.34	A
	$BC_3F_{2:3}$	NAU2002	POP12	9.67	−0.51	–	–	–
qSFC-10-1	$BC_3F_{2:3}$	CIR166a	POP11	12.79	−0.34	−0.32	0.95	A
qSFC-10-2	BC_3F_2	JESPR6	POP04	15.74	−1.60	−2.17	1.36	D
	$BC_3F_{2:3}$	JESPR6	POP35	14.29	−0.05	−0.37	7.12	H
qSFC-12-1	BC_3F_2	JESPR300	POP06	11.73	−0.37	−0.13	0.34	A
*qSFC-12-2**	BC_3F_2	DPL0866	POP11	18.96	−0.76	−0.49	0.64	A
*qSFC-12-3**	BC_3F_2	CIR293	POP32	14.58	−1.14	−0.22	0.19	A
qSFC-16-1	BC_3F_2	DPL0501	POP07	20.62	−1.82	−1.37	0.75	A
qSFC-16-2	BC_3F_2	DPL0385b	POP27	9.91	−0.57	−0.15	0.26	A
qSFC17-1	BC_3F_2	BNL2496A	POP02	12.44	−0.54	−0.08	0.15	A
*qSFC-17-2**	BC_3F_2	MUSS207	POP20	12.85	0.25	−0.40	−1.62	D
	$BC_3F_{2:4}$	MUSS207	POP20	11.50	0.17	−0.01	−0.05	A
*qSFC-19-1**	BC_3F_2	BNL3977	POP27	12.84	−0.31	−0.49	1.61	D
qSFC-19-2	BC_3F_2	DPL0140	POP08	12.83	−0.71	−0.85	1.19	D
	BC_3F_2	DPL0444	POP11	14.28	−0.68	−0.49	0.72	A
	$BC_3F_{2:3}$	DPL0140	POP11	13.56	−0.36	−0.27	0.73	A
*qSFC-19-3**	$BC_3F_{2:3}$	NAU5489	POP12	8.71	−0.35	−0.08	0.24	A
qSFC-21-1	BC_3F_2	NAU3377a	POP03	19.70	−0.62	0.01	−0.02	A
qSFC-21-2	$BC_3F_{2:3}$	CIR077	POP20	11.33	−0.25	−0.16	0.65	A
*qSFC-24-1**	BC_3F_2	TMHB1	POP01	17.29	−1.13	−0.94	0.84	A

Table 3 (continued)

QTL[a]	Generation	Locus	Family	R^2 (%)[b]	A[b]	D[b]	D/A ratio[b]	Mode of action[c]
	BC_3F_2	BNL2772a	POP01	13.99	−0.45	–	–	–
	BC_3F_2	TMHB1	POP04	12.00	−0.94	−0.88	0.93	A
*qSFC-24-2**	BC_3F_2	NAU3605	POP02	10.45	−0.56	−0.15	0.28	A
qSFC-24-3	BC_3F_2	JESPR308	POP31	21.26	−0.80	−0.46	0.58	A
qSFC-25-1	BC_3F_2	BNL4001b	POP01	19.00	−0.80	−0.49	0.61	A
qSFC-26-1	BC_3F_2	NAU3862	POP02	12.53	−0.43	0.23	−0.53	A
*qSFC-26-2**	BC_3F_2	STV122	POP16	16.62	−0.45	−0.41	0.92	A

Each row corresponds to a one-way analysis of variance for the indicated locus

[a] *Indicating that the QTL was also detected by QTLNetwork

[b]Quantitative parameters: R^2, percentage of phenotypic variation explained by the marker genotype at the corresponding marker and family. A, additive, a positive number indicates that the alleles from the *G. hirsutum* parent increase trait values; a negative number indicates that the alleles from the *G. mustelinum* parent increase trait values. *D* dominance. *D/A* ratio, overdominance effect

[c]Modes of gene action are indicated by: *A* additivity; *D* dominance; *H* overdominance. Missing values correspond to dominant marker loci

were detected for UI and SFC, namely *qUI-10-1* and *qSFC-10-2* near JESPR6; *qUI-12-2* and *qSFC-12-3* near CIR293; *qUI-12-3* and *qSFC-12-2* near DPL0866; *qUI-12-1* and *qSFC-12-1* near JESPR300; and *qUI-26-1* and *qSFC-26-1* near NAU3862. Five cases of co-location of QTLs for all three traits simultaneously were found, namely *qUHM-7-1, qUI-7-1*, and *qSFC-7-2* near MUCS616 and NAU2002; *qUHM-19-2, qUI-19-1*, and *qSFC-19-3* near NAU5489; *qUHM-24-2, qUI-24-2*, and *qSFC-24-2* near NAU3605; *qUHM-24-2, qUI-24-1*, and *qSFC-24-1* near TMHB1; and *qUHM-26-1, qUI-26-2*, and *qSFC-26-2* near STV122. The phenomenon of QTL clusters has been found in many previous studies of cotton (Chee et al. 2005b; Tang et al. 2015; Yu et al. 2013b; Zhang et al. 2011, 2015a). The nature and direction of the correlation between the traits suggest that these co-located QTLs may be conferred by pleiotropic effects of the underlying genes; however, recent evidence that cotton fiber quality QTLs may involve large numbers of co-expressed genes serving different functions offers an intriguing alternative hypothesis (Paterson et al. 2012). One way to test these hypotheses (without making transgenics) would be to develop a segregating population with recombinants in these regions of co-located QTLs (Kalladan et al. 2013; Paterson et al. 1990). In our research, QTLs for SFC showed opposite allelic effects to those of UHM or UI in most cases, with the only exception being *qUHM-19-3* and *qSFC-19-2* near DPL0140. UHM and UI always showed allelic effects in the same direction (Table 3), consistent with the correlations between them (Table 2).

Consistency of QTLs across families

Among the 218 SSR marker loci, 211 were segregating in two or more families; therefore, we performed two-way ANOVA to test for marker-trait associations and assess their consistency among families. Significant ($P<0.001$) among-family G effects were detected at 18, 16, and 16 loci for UHM, UI, and SFC, respectively (Supplementary Table 1, Table S1). For UHM, the 18 loci appeared to represent only 11 non-overlapping genomic regions on 11 different chromosomes, for which QTLs were detected near six loci in within-family analysis (*qHUM-3-1, qUHM-7-1, qUHM-21-1, qUHM-23-2, qUHM-24-1, qUHM-24-2*). For UI, the 16 loci represent 13 non-overlapping genomic regions on 13 different chromosomes, for which QTLs were detected for six loci in within-family analysis (*qUI-2-1, qUI-12-1, qUI-12-2, qUI-18-1, qUI-23-1, qUI-24-1*). For SFC, the 16 loci represent 13 non-overlapping genomic regions, for which QTLs were detected for four loci in within-family analysis (*qSFC-12-1, qSFC-12-3, qSFC-16-2, qSFC-24-1*).

A total of six loci were significant ($P<0.001$) for G×F interactions for UHM (Supplementary Table 2, Table S2). Four of these loci detected three QTLs (*qUHM-21-1, qUHM-24-2, qUHM-26-1*; Table 3, Table S2) in different segregating families, while the fifth locus (BNL2725) just missed significance ($P<0.002$) for POP16. Two loci showed G×F interactions for UI (Supplementary Table 2, Table S2). The locus BNL1151 also detected a QTL (*qUI-11-1*; Table 3, Table S2) in POP31, while the other locus (NAU915) just missed significance ($P<0.003$) for POP32.

Two loci were significant (NAU915 and NAU4045, $P < 0.001$) for G×F interactions (Supplementary Table 2, Table S2) for SFC; and both loci just missed significance ($P < 0.005$ in POP04 for NAU4045, $P < 0.002$ in POP32 for NAU915) for declaring QTLs.

Epistatic QTLs and their interactions with environments

A total of 17 epistatic QTLs with significant additive×additive (AA) effects ($P < 0.001$) were identified, with the majority (14/17) involving loci not linked to any main-effect QTLs (Table 4). Two interactions were able to be detected in a single environment and also by joint analysis: for UHM, the interaction between a region on Chr4 and a region on Chr9 was detected in both $BC_3F_{2:3}$ and joint analysis in POP10 simultaneously, with alleles from *G. mustelinum* increasing UHM; for SFC, the interaction between a region on Chr6 and a region on Chr12 was identified in both BC_3F_2 and joint analysis in POP32 simultaneously.

Discussion

The cotton genus *Gossypium* L. comprises seven tetraploid species—while *G. hirsutum* provides most of the world's cotton production, its narrow genetic diversity forms the bottle neck restricting its further improvement and leading to increase of genetic vulnerability (Paterson et al. 2004). *G. barbadense* and the five wild tetraploid cotton species [*G. tomentosum* Nutall ex Seemann, *G. mustelinum* Miers ex Watt, *G. darwinii* Watt, *G. ekmanianum* Wittmack, and *Gossypium* sp. nov.(Wendel and Grover 2015)] provide resources that can be used to improve *G. hirsutum* traits, including fiber quality. Many studies have been performed to exploit elite genes from *G. barbadense* to improve fiber quality in *G. hirsutum* (Chee et al. 2005a, b; Draye et al. 2005; Lacape et al. 2010; Rong et al. 2004; Said et al. 2015; Shi et al. 2015; Wang et al. 2011, 2013; Yu et al. 2013a, b, 2014b). Exploratory efforts have also been made to construct genetic linkage maps and/or identify fiber quality QTLs from *G. tomentosum* (Hou et al. 2013; Khan et al. 2016; Waghmare et al. 2005; Zhang et al. 2011) and *G. darwinii* (Chen et al. 2015; Wang et al. 2012).

The results of this study build on a rich and growing body of evidence that DNA marker-assisted approaches offer the potential to extract from wild tetraploid cottons alleles of value in improvement of elite cultivated cottons. *G. mustelinum*, the basal species in the tetraploid *Gossypium* clade, is a wild tetraploid cotton endemic to the semi-arid region of northeastern Brazil and is sexually compatible with cultivated cotton (*G. hirsutum* or *G. barbadense*). It is not known to have yet been bred or commercially exploited (Alves et al. 2013; Borém et al. 2003), or subjected to domestic use by indigenous people in its range that may have contributed to the improvement of important traits such as fiber quality (Alves et al. 2013).

In this study, AB-QTL analysis was performed based on SSR markers and phenotypic data from three generations of BC_3-derived families. Phenotypic evaluation of the mapping population showed transgressive segregation in the progenies for all fiber length traits in each of the three generations, indicating the presence of both positive and negative alleles for each trait in each parent (Supplementary Fig. 1; Table 1). Since *G. mustelinum* has inferior fiber length traits relative to the *G. hirsutum* parent, it is not a surprise to find that most BC_3 families (15 of 21 in BC_3F_2, eight of 12 in $BC_3F_{2:3}$, and nine of 12 in $BC_3F_{2:4}$) have lower mean UHM values than the recurrent *G. hirsutum* parent, PD94042. However, it is interesting that for UI and SFC, most BC_3F_2 families (15 of 21 for UI, and 16 of 21 for SFC) outperformed the recurrent parent (Supplementary Fig. 1). Swamy et al. (2012) reported transgressive segregation for most grain quality traits of rice with 75% of families showing at least 5% increase over the cultivated parent for several traits. Wickneswari et al. (2012) also found that 62.8 and 74.8% of BC_2F_2 families outperformed the recurrent parent for grain yield and 1000 grain weight. In our research, even though *G. mustelinum* has inferior fiber length, it contributed positively to UI and SFC in most BC_3F_2 families and also to UHM in some families. In these families, many individual plants have better fiber length traits than the recurrent parent (Table 1). The transgressive segregation suggests that many new gene combinations formed by interspecific hybridization are desirable, although some are clearly not; the accumulation and interaction of superior alleles from the two divergent parents may have played an important role in phenotype expression. Such 'positive' transgression is of particular importance in view of the genetic bottlenecks that have constrained variation in the elite Upland cotton gene pool (Paterson et al. 2004).

Advanced backcross mapping populations have more uniform genetic background than early generation crosses, which increases the ability to detect even QTLs with small effects (Nagata et al. 2015; Tanksley and Nelson 1996). Advanced backcross populations have been created in many crops and made noteworthy contributions in revealing favorable alleles from wild species (Wang and Chee 2010), such as tomato (Fulop et al. 2016; Kinkade and Foolad 2013), barley (Haas et al. 2016; Kalladan et al. 2013), rice (Kim et al. 2015; Nagata et al. 2015), maize (Trachsel et al. 2016), peanut (Burow et al. 2014), and wheat (Naz et al. 2015). Some of these QTLs were further fine-mapped and cloned, e.g., six genes controlling fruit weight and shape

Table 4 Estimated epistasis and epistasis × environment interaction effects of QTLs for fiber length traits

Trait	Env.[a]	Family	QTL_i[b]	Chrosome	Interval_i	QTL_j[b]	Chrosome	Interval_j	AA[c]	*P* Value	h^2 (aa) (%)[d]	h^2 (aae) (%)[e]
UHM	BC_3F_2	POP34	–	Chr19	DPL0444-BNL3903	–	Chr11	BNL1681-BNL1151	−1.64	0.000002	16.24	
	BC_3F_2	POP10	–	Chr22	CIR048-BNL530a	*qUHM-23-2*	Chr23	BNL3383-BNL3511	0.76	0.000006	13.37	
	$BC_3F_{2:3}$	POP10	–	Chr4	DPL0196a-MUSB1050	–	Chr9	DPL0507-BNL3031b	−0.84	0.000028	13.56	
	Joint	POP10	–	Chr4	DPL0196a-MUSB1050	–	Chr9	DPL0507-BNL3031b	−0.55	0.000001	6.47	0.51
	Joint	POP11	–	Chr1	MUSS085-MUSS161	–	Chr4	DPL0196a-MUSB1050	−0.83	0.000000	7.94	1.37
	Joint	POP10		Chr3	DPL0354-DPL0605	–	Chr19	BNL3811-BNL3977	−0.65	0.000000	7.23	1.23
	Joint	POP34	–	Chr5	CIR102-DPL0241		Chr10	BNL1161-CIR166a	1.01	0.000000	6.35	0.37
	Joint	POP34	–	Chr11	BNL4094-BNL1681	–	Chr18	STS1155b-NAU2488	−0.67	0.000002	5.14	0.60
	Joint	POP15	–	Chr15	BNL1350-BNL2646		Chr23	DPL0378-DPL0262a	−0.40	0.000072	3.17	0.37
	Joint	POP17	–	Chr19	NAU3205-BNL3535a	–	Chr26	BNL3816-BNL341	0.48	0.000009	4.22	0.95
	Joint	POP35	*qUHM-19-3*	Chr19	DPL0140-BNL2715		Chr23	DPL0378-DPL0262a	0.66	0.000029	3.34	0.44
UI	BC_3F_2	POP27	–	Chr3	DPL0605-BNL3441		Chr12	CIR293-NAU1151	0.91	0.000029	10.11	
	$BC_3F_{2:3}$	POP34		Chr10	CIR171-JESPR6		Chr11	BNL1681-BNL1151	−0.83	0.000126	10.71	
	$BC_3F_{2:4}$	POP31	–	Chr3	DPL0605-BNL3441		Chr19	NAU5489-BNL632	0.55	0.000289	14.40	
	Joint	POP34		Chr5	DPL0241-BNL3029	–	Chr11	BNL1681-BNL1151	0.48	0.000025	1.42	0.90
SFC	BC_3F_2	POP32	*qSFC-6-1*	Chr6	NAU5433-DPL0183a	*qSFC-12-3*	Chr12	NAU915-CIR293	−1.73	0.000011	14.33	
	Joint	POP32	*qSFC-6-1*	Chr6	NAU5433-DPL0183a	*qSFC-12-3*	Chr12	CIR293-NAU2640b	−0.94	0.000000	0.71	1.26
	BC_3F_2	POP03		Chr12	NAU2640B-JESPR300		Chr18	NAU2488-MUSS603	−0.81	0.000258	22.15	
	Joint	POP11	–	Chr19	CIR212-BNL285		Chr7	DPL0234-MUSS006	−0.30	0.000088	3.06	3.71

[a]Joint: results obtained based on combined data of the BC_3F_2, $BC_3F_{2:3}$, and $BC_3F_{2:4}$ generations

[b]QTL with main effect of locus i or j detected in one-way analysis of variance in the same family

[c]Epistatic effects of the additive × additive interaction. A positive number indicates that the *G. hirsutum* alleles increase trait values; a negative number indicates that the *G. mustelinum* alleles increase trait values

[d]Phenotypic variance explained by additive × additive interaction effects

[e]Phenotypic variance explained by AA by environment effect

variation in tomato played important roles in tomato breeding, and may also aid fundamental and applied research in other plants (van der Knaap et al. 2014). However, linkage of QTL alleles in repulsion may decrease the ability to detect QTLs and result in underestimation of the additive effect when the two QTLs are segregating (Kroymann and Mitchell-Olds 2005; Nagata et al. 2015). In addition, problems may exist in QTL mapping for families only in the BC_3F_2 generation, since genotype by environment interactions affect phenotype evaluation (Table 4)—phenotype of only a single individual was measured for each genotype in all BC_3F_2 families, which might also lead to escape of some QTLs from detection. By also evaluating $BC_3F_{2:3}$ and $BC_3F_{2:4}$ generations for 12 of the 21 families with enough seeds, using multiple individuals of each genotype, QTLs mapped in the BC_3F_2 generation can be validated, also detecting additional ones missed in the BC_3F_2.

From the three generations studied, one-way analysis of variance detected 19, 20, and 26 non-overlapping QTLs with PVE of 16.06%, 14.66%, and 13.77% on average for UHM, UI, and SFC, respectively; twelve of the QTLs explained more than 20% of PV (Table 3). As expected, *G. hirsutum* alleles increased UHM and UI and reduced SFC for most QTLs, consistent with the parental fiber phenotypes. The effects of many QTLs showed good reproducibility, with 23 detected at least twice near adjacent markers in the same family or near the same markers across different families/generations. In addition, 32 of the 65 QTLs were detected in both one-way variance analyses and the MCIM method of QTLNetwork (Table 3). Although the *G. mustelinum* parent does not produce spinnable fiber, *G. mustelinum* alleles increased UHM for five of the 19 QTLs, also increasing UI for *qUI-14-1* and decreasing SFC for *qSFC-17-2*, which can be detected in two generations in POP20—these QTLs are of great interest to be further exploited by transfer into Upland cotton. Efforts are now underway to construct near-isogenic introgression lines that will allow these *G. mustelinum* alleles to be more readily accessible in breeding programs for improving fiber quality.

Using data accumulated over three generations, epistatic QTLs were detected in addition to main-effect QTLs for fiber length (Table 4). Both main-effect and epistatic QTLs play important genetic roles in controlling cotton fiber length, with environment also playing a role but to a lesser degree than additive and epistatic effects (Table 4). It is interesting that most of the epistatic interactions (14/17) were detected between genetic background loci not linked to any QTL (Table 4), indicating the complexity of fiber length inheritance. Naturally, introgressing and testing epistatic alleles involves at least twice the difficulty associated with utilizing main-effect alleles from *G. mustelinum* (Chee et al. 2005b; Tang et al. 2015; Yu et al. 2013b; Zhang et al. 2011, 2015a). More than one *G. mustelinum* BC_3 family often segregated for the same chromosomal segment in our research, making it possible to investigate the effects of genetic background on introgressed chromatin. Significant ($P < 0.001$) among-family G effects were detected at 18, 16, and 16 loci for UHM, UI, and SFC, respectively, with several of these loci revealing QTLs in within-family analysis (Supplementary Table 1, Table S1). Some among-family G effects showed good reproducibility, with six of 18 for UHM, two of 16 for UI, and one of 16 for SFC detected in different generations (Table S1). In addition, six, two, and two QTL regions for UHM, UI, and SFC present in multiple families showed G×F interactions, indicating that the introgressed region may have a strong positive effect on these traits in one family but a lesser effect in other families (Table 3, Table S2). Similar interactions between QTL and genetic backgrounds were also observed in Upland cotton populations introgressed with *G. barbadense* (Chee et al. 2005a, b; Draye et al. 2005) and *G. tomentosum* chromosome segments (Zhang et al. 2011).

This study also add further to widespread reports of the large contribution to the genetic control of fiber quality traits of the tetraploid D-subgenome, derived from a diploid progenitor that does not produce spinnable fiber (Chee et al. 2005b; Jiang et al. 1998; Zhang et al. 2008, 2011). In the present study, among the 65 QTLs that affect fiber length properties, the D-subgenome accounted for 39, much more than the 26 in the A-subgenome. The use of interspecific gene combinations appears to offer an important means by which to improve fiber length traits in Upland cotton. Although barriers to gene introgression from *G. mustelinum* to *G. hirsutum* may exist, the availability of DNA markers linked to *G. mustelinum* QTLs identified in this research could assist breeders in transferring and maintaining these traits during cultivar development. Further, since many QTLs from *G. mustelinum* are now in a near-isogenic state in the AB-QTL population, the phenotypic effect evaluated for each QTL measured here is likely to better predict its eventual effect when transferred to other cultivated backgrounds.

Author contribution statement AHP conceived and designed the experiments; PWC and OLM oversaw crossing, phenotyping and genotyping, which was performed by BW, ZZ, ZZ, and ELL; DJ oversaw fiber analysis; BW, XD, and ML analyzed the data; BW, AHP, and PWC wrote the paper.

Acknowledgements We appreciate financial help from the National Science Foundation (PFI award IIP 0917856; AIR award IIP 1127755), Cotton Incorporated, and Key Research and Development Project of Jiangsu Province, China (Modern Agriculture, BE2015353). We thank Jennifer McCurdy for help in field work and Cornelia Lemke for technical assistance in the lab.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Alves M, Barroso P, Ciampi A, Hoffmann L, Azevedo V, Cavalcante U (2013) Diversity and genetic structure among subpopulations of *Gossypium mustelinum* (Malvaceae). Genet Mol Res 12:597–609

Borém A, Freire EC, Penna JCV, Barroso PAV (2003) Considerations about cotton gene escape in Brazil: a review. Crop Breed Appl Biotechnol 3:315–332

Burow MD, Starr JL, Park CH, Simpson CE, Paterson AH (2014) Introgression of homeologous quantitative trait loci (QTLs) for resistance to the root-knot nematode [*Meloidogyne arenaria* (Neal) Chitwood] in an advanced backcross-QTL population of peanut (*Arachis hypogaea* L.). Mol Breed 34:393–406

Chee P, Draye X, Jiang CX, Decanini L, Delmonte TA, Bredhauer R, Smith CW, Paterson AH (2005a) Molecular dissection of interspecific variation between *Gossypium hirsutum* and *Gossypium 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 TA, Bredhauer R, Smith CW, Paterson AH (2005b) Molecular dissection of phenotypic variation between *Gossypium hirsutum* and *Gossypium barbadense* (cotton) by a backcross-self approach: III. Fiber length. Theor Appl Genet 111:772–781

Chen H, Khan MKR, Zhou Z, Wang X, Cai X, Ilyas MK, Wang C, Wang Y, Li Y, Liu F, Wang K (2015) A high-density SSR genetic map constructed from a F_2 population of *Gossypium hirsutum* and *Gossypium darwinii*. Gene 574:273–286

Draye X, Chee P, Jiang CX, Decanini L, Delmonte TA, Bredhauer R, Smith CW, Paterson AH (2005) Molecular dissection of interspecific variation between *Gossypium hirsutum* and *G. barbadense* (cotton) by a backcross-self approach: II. Fiber fineness. Theor Appl Genet 111:764–771

Fulop D, Ranjan A, Ofner I, Covington MF, Chitwood DH, West D, Ichihashi Y, Headland L, Zamir D, Maloof JN, Sinha NR (2016) A new advanced backcross tomato population enables high resolution leaf QTL mapping and gene identification. G3-Genes Genom Genet 6:3169–3184

Haas M, Menke J, Chao S, Steffenson BJ (2016) Mapping quantitative trait loci conferring resistance to a widely virulent isolate of *Cochliobolus sativus* in wild barley accession PI 466423. Theor Appl Genet 129:1–12

Hou MY, Cai CP, Zhang SW, Guo WZ, Zhang TZ, Zhou BL (2013) Construction of microsatellite-based linkage map and mapping of nectarilessness and hairiness genes in *Gossypium tomentosum*. J Genet 92:445–459

Jiang CX, Wright RJ, El-Zik KM, Paterson AH (1998) Polyploid formation created unique avenues for response to selection in *Gossypium* (cotton). Proc Natl Acad Sci USA 95:4419–4424

Kalladan R, Worch S, Rolletschek H, Harshavardhan VT, Kuntze L, Seiler C, Sreenivasulu N, Roder MS (2013) Identification of quantitative trait loci contributing to yield and seed quality parameters under terminal drought in barley advanced backcross lines. Mol Breed 32:71–90

Khan MKR, Chen H, Zhou Z, Ilyas MK, Wang X, Cai X, Wang C, Liu F, Wang K (2016) Genome wide SSR high density genetic map construction from an interspecific cross of *Gossypium hirsutum* × *Gossypium tomentosum*. Front Plant Sci 7:436

Kim SM, Suh JP, Qin Y, Noh TH, Reinke RF, Jena KK (2015) Identification and fine-mapping of a new resistance gene, *Xa40*, conferring resistance to bacterial blight races in rice (*Oryza sativa* L.). Theor Appl Genet 128:1933–1943

Kinkade MP, Foolad MR (2013) Validation and fine mapping of *lyc12.1*, a QTL for increased tomato fruit lycopene content. Theor Appl Genet 126:2163–2175

Kroymann J, Mitchell-Olds T (2005) Epistasis and balanced polymorphism influencing complex trait variation. Nature 435:95–98

Kuang XQ, Yu CW (2015) Generating cotton fiber length probability density function from fiber length parameters by finite mixture model. J Text I 106:655–662

Lacape JM et al (2010) Meta-analysis of cotton fiber quality QTLs across diverse environments in a *Gossypium hirsutum* x *G. barbadense* RIL population. BMC Plant Biol 10:132

Lawrence CA (2003) Fundamentals of spun yarn technology. CRC Press, Boca Raton

McCouch S, Cho Y, Yano M, Paul E, Blinstrub M, Morishima H, Kinoshita T (1997) Report on QTL nomenclature. Rice Genet Newsl 14:11

Nagata K et al (2015) Advanced backcross QTL analysis reveals complicated genetic control of rice grain shape in a *japonica* x *indica* cross. Breed Sci 65:308–318

Naz AA, Klaus M, Pillen K, Leon J (2015) Genetic analysis and detection of new QTL alleles for *Septoria tritici* blotch resistance using two advanced backcross wheat populations. Plant Breed 134:514–519

Paterson AH, Deverna JW, Lanini B, Tanksley SD (1990) Fine mapping of quantitative trait loci using selected overlapping recombinant chromosomes, in an interspecies cross of tomato. Genetics 124:735–742

Paterson AH et al (2004) Reducing the genetic vulnerability of cotton. Crop Sci 44:1900–1901

Paterson AH et al (2012) Repeated polyploidization of *Gossypium* genomes and the evolution of spinnable cotton fibres. Nature 492:423–427

Rong JK et al (2004) A 3347-locus genetic recombination map of sequence-tagged sites reveals features of genome organization, transmission and evolution of cotton (*Gossypium*). Genetics 166:389–417

Said JI, Song M, Wang H, Lin Z, Zhang X, Fang DD, Zhang J (2015) A comparative meta-analysis of QTL between intraspecific *Gossypium hirsutum* and interspecific *G. hirsutum* × *G. barbadense* populations. Mol Genet Genomics 290:1003–1025

SAS Institute (1999) SAS/STAT user's guide, version 8, vol 2. SAS Institute, Cary

Self SG, Liang K-Y (1987) Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. J Am Stat Assoc 82:605–610

Shi YZ et al (2015) Constructing a high-density linkage map for *Gossypium hirsutum* x *Gossypium barbadense* and identifying QTLs for lint percentage. J Integr Plant Biol 57:450–467

Swamy BPM, Kaladhar K, Rani NS, Prasad GSV, Viraktamath BC, Reddy GA, Sarla N (2012) QTL analysis for grain quality traits in 2 BC_2F_2 populations derived from crosses between *Oryza sativa* cv Swarna and 2 accessions of *O. nivara*. J Hered 103:442–452

Tang SY et al (2015) Construction of genetic map and QTL analysis of fiber quality traits for Upland cotton (*Gossypium hirsutum* L.). Euphytica 201:195–213

Tanksley S, Nelson J (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

Testore F, Minero G (1988) A study of the fundamental parameters of some fancy yarns. J Text I 79:606–619

Trachsel S, Sun D, SanVicente FM, Zheng H, Atlin GN, Suarez EA, Babu R, Zhang X (2016) Identification of QTL for early vigor and stay-green conferring tolerance to drought in two connected advanced backcross populations in tropical maize (*Zea mays* L.). PLoS ONE 11(3):e0149636

van der Knaap E et al (2014) What lies beyond the eye: the molecular mechanisms regulating tomato fruit weight and shape. Front Plant Sci 5:227

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

Wang B, Chee PW (2010) Application of advanced backcross quantitative trait locus (QTL) analysis in crop improvement. J Plant Breed Crop Sci 2:221–232

Wang FR, Gong YC, Zhang CY, Liu GD, Wang LM, Xu ZZ, Zhang J (2011) 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 B, Nie YC, Lin ZX, Zhang XL, Liu JJ, Bai J (2012) Molecular diversity, genomic constitution, and QTL mapping of fiber quality by mapped SSRs in introgression lines derived from *Gossypium hirsutum* x *G. darwinii* Watt. Theor Appl Genet 125:1263–1274

Wang XQ, Yu Y, Li W, Guo HL, Lin ZX, Zhang XL (2013) Association analysis of yield and fiber quality traits in *Gossypium barbadense* with SSRs and SRAPs. Genet Mol Res 12:3353–3362

Wang B, Draye X, Zhang Z, Zhuang Z, May OL, Paterson AH, Chee PW (2016a) Advanced backcross QTL analysis of fiber elongation in a cross between *Gossypium hirsutum* and *G. mustelinum*. Crop Sci 56:1760–1768

Wang B, Liu L, Zhang D, Zhuang Z, Guo H, Qiao X, Wei L, Rong J, May OL, Paterson AH, Chee PW (2016b) A genetic map between *Gossypium hirsutum* and the Brazilian endemic *G. mustelinum* and its application to QTL mapping. G3-Genes Genomics Genet 6:1673–1685

Wendel JF, Grover CE (2015) Taxonomy and evolution of the cotton genus, *Gossypium*. In: Fang DD, Percy RG (ed) Cotton, 2nd edn. American Society of Agronomy, Inc., Crop Science Society of America, Inc., and Soil Science Society of America, Inc, Madison.

Wickneswari R, Bhuiyan MAR, Kalluvettankuzhy S, Lim LS, Thomson MJ, Narimah MK, Abdullah MZ (2012) Identification and validation of quantitative trait loci for agronomic traits in advanced backcross breeding lines derived from *Oryza rufipogon* x *Oryza sativa* cultivar MR219. Plant Mol Biol Rep 30:929–939

Yang J, Hu CC, Hu H, Yu RD, Xia Z, Ye XZ, Zhu J (2008) QTLNetwork: mapping and visualizing genetic architecture of complex traits in experimental populations. Bioinformatics 24:721–723

Yu JW, Yu SX, Gore M, Wu M, Zhai HH, Li XL, Fan SL, Song MZ, Zhang JF (2013a) Identification of quantitative trait loci across interspecific F_2, $F_{2:3}$ and testcross populations for agronomic and fiber traits in tetraploid cotton. Euphytica 191:375–389

Yu JW, Zhang K, Li SY, Yu SX, Zhai HH, Wu M, Li XL, Fan SL, Song MZ, Yang DG, Li YH, Zhang JF (2013b) Mapping quantitative trait loci for lint yield and fiber quality across environments in a *Gossypium hirsutum* x *Gossypium barbadense* backcross inbred line population. Theor Appl Genet 126:275–287

Yu J, Jung S, Cheng CH, Ficklin SP, Lee T, Zheng P, Jones D, Percy RG, Main D (2014a) CottonGen: a genomics, genetics and breeding database for cotton research. Nucleic Acids Res 42:D1229–D1236

Yu JZ, Ulloa M, Hoffman SM, Kohel RJ, Pepper AE, Fang DD, Percy RG, Burke JJ (2014b) Mapping genomic loci for cotton plant architecture, yield components, and fiber properties in an interspecific (*Gossypium hirsutum* L. x *G. barbadense* L.) RIL population. Mol Genet Genomics 289:1347–1367

Zhang HB, Li Y, Wang B, Chee PW (2008) Recent advances in cotton genomics. Int J Plant Genomics 2008:742304

Zhang ZS, Rong JK, Waghmare VN, Chee PW, May OL, Wright RJ, Gannaway JR, Paterson AH (2011) QTL alleles for improved fiber quality from a wild Hawaiian cotton, *Gossypium tomentosum*. Theor Appl Genet 123:1075–1088

Zhang JF, Yu JW, Pei WF, Li XL, Said J, Song MZ, Sanogo S (2015a) Genetic analysis of Verticillium wilt resistance in a backcross inbred line population and a meta-analysis of quantitative trait loci for disease resistance in cotton. BMC Genomics 16:577

Zhang TZ, Hu Y, Jiang WK, Fang L, Guan XY, Chen JD, Zhang JB, Saski CA, Scheffler BE, Stelly DM, Hulse-Kemp AM, Wan Q, Liu BL, Liu CX, Wang S, Pan MQ, Wang YK, Wang DW, Ye WX, Chang LJ, Zhang WP, Song QX, Kirkbride RC, Chen XY, Dennis E, Llewellyn DJ, Peterson DG, Thaxton P, Jones DC, Wang Q, Xu XY, Zhang H, Wu HT, Zhou L, Mei GF, Chen SQ, Tian Y, Xiang D, Li XH, Ding J, Zuo QY, Tao LN, Liu YC, Li J, Lin Y, Hui YY, Cao ZS, Cai CP, Zhu XF, Jiang Z, Zhou BL, Guo WZ, Li RQ, Chen ZJ (2015b) Sequencing of allotetraploid cotton (*Gossypium hirsutum* L. acc. TM-1) provides a resource for fiber improvement. Nat Biotechnol 33:531–537