Advanced Backcross QTL Analysis of Fiber Strength and Fineness in a Cross between Gossypium hirsutum and G. mustelinum

Baohua Wang1,2,3, Zhimin Zhuang1,3, Zhengsheng Zhang2,4, Xavier Draye5, Lan-Shuan Shuang2, Tariq Shehzad2, Edward L. Lubbers3, Don Jones6, O. Lloyd May3, Andrew H. Paterson2* and Peng W. Chee3*

1 School of Life Sciences, Nantong University, Nantong, China, 2 Plant Genome Mapping Laboratory, University of Georgia, Athens, GA, United States, 3 Department of Crop and Soil Sciences, University of Georgia, Tifton, GA, United States, 4 Engineering Research Center of South Upland Agriculture, Ministry of Education, Southwest University, Chongqing, China, 5 Earth and Life Institute, Université catholique de Louvain, Louvain-la-Neuve, Belgium, 6 Agricultural Research Division, Cotton Incorporated, Cary, NC, United States

The molecular genetic basis of cotton fiber strength and fineness in crosses between Gossypium mustelinum and Gossypium hirsutum (Upland cotton) was dissected using 21 BC3F2 and 12 corresponding BC3F2:3 and BC3F2:4 families. The BC3F2 families were genotyped with simple sequence repeat markers from a G. hirsutum by G. mustelinum linkage map, and the three generations of BC3-derived families were phenotyped for fiber strength (STR) and fineness (Micronaire, MIC). A total of 42 quantitative trait loci (QTLs) were identified through one-way analysis of variance, including 15 QTLs for STR and 27 for MIC, with the percentage of variance explained by individual loci averaging 13.86 and 14.06%, respectively. Eighteen of the 42 QTLs were detected at least twice near the same markers in different generations/families or near linked markers in the same family, and 28 of the 42 QTLs were identified in both mixed model-based composite interval mapping and one-way variance analyses. Alleles from G. mustelinum increased STR for eight of 15 and reduced MIC for 15 of 27 QTLs. Significant among-family genotypic effects (P < 0.001) were detected in 13 and 10 loci for STR and MIC respectively, and five loci showed significant (P < 0.001) genotype × family interaction for MIC. These results support the hypothesis that fiber quality improvement for Upland cotton could be realized by introgressing G. mustelinum alleles although complexities due to the different effects of genetic background on introgressed chromatin might be faced. Building on prior work with G. barbadense, G. tomentosum, and G. darwinii, QTL mapping involving introgression of G. mustelinum alleles offers new allelic variation to Upland cotton germplasm.

Keywords: genetic diversity, introgression, marker-assisted breeding, fiber strength, fiber fineness

INTRODUCTION

This is the fourth report describing the interspecific G. mustelinum by G. hirsutum (Upland cotton) genetic map and mapping and introgressing quantitative trait loci (QTLs) of fiber quality traits from G. mustelinum into Upland cotton. In three previous papers, we reported the interspecific G. hirsutum by G. mustelinum genetic linkage map (Wang et al., 2016b), and described 24 QTLs...
for fiber elongation (Wang et al., 2016a,b) and 65 QTLs for fiber length traits (Wang et al., 2017). Some alleles from G. mustelinum improved fiber properties, demonstrating their potential value for improving fiber quality in Upland cotton breeding.

Here interspecific QTL mapping and introgression were reported for two important fiber quality traits, fiber strength (STR) and fineness measured in Micronaire (MIC). STR is determined as the necessary force breaking a fiber “beard.” Following the measurement of length by the Uster High Volume Instrument (HVI), the measurement of strength is performed on the same fiber bead with two sets of jaws clamping the bead by using a gauge length of 3.175 mm. The breaking force is measured directly and normalized with an assessment of the fiber mass from the optical sensor (in combination with the MIC value) to give the strength in cN/tex (Naylor et al., 2014). Fiber tenacity usually affects yarn tenacity more than any other properties of fiber at optimum yarn twist, and when fiber strength increases one cN/tex, yarn strength will increase about 0.5 cN/tex or even more. Generally bundle tenacity higher than 30 cN/tex is considered desirable (Estur and Knappe, 2007). Usually STR should be increased in cotton breeding programs.

MIC is a measure of the air permeability of compressed cotton fibers, which reflects both maturity in the development degree of cell walls and fiber fineness measured in linear density. A constant cotton fiber mass is compressed into a space with known volume. This compressed sample is used to measure air permeability, which is converted to appropriate numbers denoting MIC values. Generally MIC readings of 3.7–4.2 are premium, 3.5–3.6 or 4.3–4.9 are base, and 3.4-and-under or 5.0-and-higher are substandard and may result in a discounted price to the producer (http://www.cottoninc.com/fiber/quality/Classification-Of-Cotton/Classing-booklet.pdf). When the measurement comes in too low, the cotton is more susceptible to entangling around debris, which means too much of the good fiber will also be lost. When it is too high, it also causes problems since a coarser fiber negatively affects the spinning process, as well as overall quality is undesirable from the aspect of yarn evenness and spinning (Montalvo, 2005). The fineness determines how many fibers are present in the cross-section of a yarn of given thickness. More fibers available in the cross section of yarn generally results in stronger yarn, usually produced with finer fibers (Estur and Knappe, 2007). Additional fibers in the cross-section provide not only additional strength but also better evenness in the yarn. Many Upland cotton varieties have the problem of high MIC values currently, so it is necessary to breed varieties with comparatively low MIC or finer fiber.

This research aims to map QTLs for STR and MIC in a set of advanced-backcross G. hirsutum × G. mustelinum population. The QTLs mapped in this research will enhance our understanding of the molecular genetic basis of cotton fiber quality. They will also benefit cotton molecular breeding to improve STR and MIC with G. mustelinum alleles and to ascertain the specific genetic basis of these important traits.

MATERIALS AND METHODS

Population Development and Field Evaluation

Three generations of interspecific advanced-backcross populations, namely 21 BC3F2 families and 12 BC3F2:3 and BC3F2:4 families were developed as follows: a G. hirsutum acc., PD94042 was crossed with G. mustelinum (AD4-8), and then F1 plants were independently backcrossed to the G. hirsutum parent for three cycles; a total of 21 lineages that produced BC3F1 plants were self-pollinated and generated 21 BC3F2 families with size ranging from 127 to 160 plants (totally 3,203 BC3F2 progenies; Wang et al., 2016a); in addition, 12 BC3F2:3 and BC3F2:4 families with size ranging from 130 to 160 lines (totally 1,826 lines) were developed for 12 of the 21 BC3F2 families with enough seeds (Table S1). The 21 BC3F2 families were planted in 2006; completely randomized designs were used for the 12 BC3F2:3 and BC3F2:4 families with two replicate plots in two years (2008 and 2009) in Tifton, Georgia. All cultural practices were performed as described in Wang et al. (2016a). For the BC3F2 generation, seed-cotton of all bolls was hand-harvested for each plant; for the BC3F2:3 and BC3F2:4 generations, seed-cotton was hand-picked from two random replicate plots as two samples. Seed-cottons were ginned on a saw gin, and then STR and MIC were tested by the Cotton Incorporated Textile Services Laboratory (Cotton Incorporated, Cary, N.C.) by using the Uster High Volume Instrument (HVI; Zellweger-Uster, Knoxville, Tenn.).

Genotyping and Data Analysis

The BC3F1 plants were genotyped with 218 SSR markers selected for even representation of our interspecific G. mustelinum by G. hirsutum map which comprised 1,055 loci (Wang et al., 2016b), constructed with an F2 population of the same two parents. The markers with introgression from G. mustelinum in the BC3F1:8 were then screened in the corresponding BC3F2 families for genotyping (Wang et al., 2016a). An average of 58 markers were used to genotype each BC3F2 family, with marker numbers ranging from 47 to 81, and the size of BC3F2 families ranged from 127 to 160. This set of genotype data based on each BC3F2 individual also constitutes the genotype of the corresponding BC3F2:3 and BC3F2:4 lines.

For every marker locus that segregated within BC3F2 families, one-way variance analyses were used to test associations between phenotypes and marker genotypes for statistical significance. The GLM procedure in the SAS ver.8 software package (SAS Institute, 1999) was used to carry out the analyses, and the significance threshold was set at $P < 0.001$ for F-test. The gene action modes (additive, a and dominant, d) for individual QTLs were evaluated with their significance levels estimated as described by Paterson et al. (1990). The dominance/additive (d/a) ratio of 3 was used as the threshold to determine whether the QTLs were over- or under-dominant (Chee et al., 2005a).

QTLs were also detected with the software QTLNetwork V2.1 (Yang et al., 2008) for STR and MIC in each of the BC3-derived families, so as to map epistatic QTLs and also help confirm the reliability of the QTLs identified by one-way variance analyses. The critical F value calculated based on 1,000 permutation tests...
Wang et al. QTL Mapping of Fiber Strength and Fineness was used in the mixed model-based composite interval mapping (MCIM) method, with walk speed and window size set at 1 and 10 cM, respectively. A putative main-effect or epistatic QTL was claimed with the significance threshold set at $P = 0.001$. Considering environmental effects, QTLs were also mapped in joint analysis for the 12 families that were grown in three different environments/generations ($BC_3F_2$, $BC_3F_2:3$, and $BC_3F_2:4$) with the software QTLNetwork V2.1 (Yang et al., 2008). QTLs sharing a common marker between the two methods (QTLNetwork 2.1 and one-way variance analyses) were regarded as the same QTLs. QTL Nomenclature of QTLs was performed as previously described by McCouch et al. (1997); the QTL name began with a “q,” representing a QTL, followed by an abbreviation of the trait name, the chromosome name, and consecutive numbers indicating the QTL number of the same trait on the same chromosome (Wang et al., 2017).

For the loci that segregated in two or more families, two-way mixed model variance analyses were also applied, using the MIXED procedure of the SAS ver. 8 package, including genotype ($G$) as a fixed factor and family ($F$) and genotype × family ($G \times F$) interaction as random factors, and the residual maximum likelihood (REML) method was used to estimate model parameters. Marker-trait association (genotype factor) was evaluated with an $F$ statistic using a general Satterthwaite approximation for the denominator degrees of freedom (SAS Institute, 1999). A likelihood-ratio ($\chi^2$) test was carried out for the $G \times F$ interaction (Self and Liang, 1987; Chee et al., 2005a). $P < 0.001$ was set as the significance level for both $G$ and $G \times F$ effects.

RESULTS

Phenotypic Distribution and Correlations
The distributions of STR and MIC of the $BC_3$ progenies are shown in Figure 1 and Table 1. Both traits expressed significant segregation in the three generations, and their distribution ranges were wider in $BC_3F_2$ than $BC_3F_2:3$ and $BC_3F_2:4$ (Table 1). Although $G. mustelinum$, the wild parent cannot produce spinnable fiber, many $BC_3$ progenies show higher STR and lower MIC (usually lower MIC is preferred) than the cultivated parent.

![Figure 1](image1.png)  
**FIGURE 1** | Phenotypic distribution of fiber strength (STR) and fineness (Micronaire, MIC) demonstrated by family mean deviation from recurrent Gossypium hirsutum (GH) parent.
Table 1: Summary statistics of fiber strength and fineness measured on the G. hirsutum parent and BC3 progenies.

| Generation | Trait                     | Progeny | G. hirsutum |
|------------|---------------------------|---------|-------------|
|            | Range                     | Mean ± SD | CV (%)     | Parent |
| BC3F2      | Fiber strength (STR, cN/tex) | 17.5–35.3 | 27.1 ± 2.27 | 8.4 | 28.0 |
|            | Micronaire (MIC)          | 2.0–6.6 | 4.3 ± 0.74 | 17.2 | 4.6 |
| BC3F2,3    | Fiber strength (STR, cN/tex) | 20.2–41.9 | 30.5 ± 2.35 | 7.7 | 29.9 |
|            | Micronaire (MIC)          | 2.4–5.99 | 4.5 ± 0.54 | 12.0 | 4.7 |
| BC3F2,4    | Fiber strength (STR, cN/tex) | 25.4–37.7 | 31.0 ± 1.63 | 5.3 | 30.6 |
|            | Micronaire (MIC)          | 3.3–5.6 | 4.6 ± 0.38 | 8.3 | 4.6 |

Main-Effect QTLs Detected for Each Trait

By assuming that each block of linked markers showing significant marker-trait association (P < 0.001) within a family denoted a single QTL, a total of 42 non-overlapping QTLs were identified in the three generations of BC3-derived families (Figure 2 and Table 2). These QTLs were mapped to 20 chromosomes, with 22 located on 12 A-subgenome chromosomes and 20 located on eight D-subgenome chromosomes. One or more QTLs for STR and MIC were identified in 13 of the 21 families, with a maximum of five QTLs in each of three families (POP12, POP15, and POP32). Eighteen of the 42 QTLs could be identified by QTLNetwork. The detailed QTL information is listed in Figure 2 and Table 2.

QTLs for MIC

A total of 27 non-overlapping QTLs for MIC were identified in 15 chromosomes with 15 located on nine A-subgenome chromosomes and 12 located on six D-subgenome chromosomes (Figure 2 and Table 2). Eleven QTLs were found at least twice near the same markers in different generations/families or near linked markers in the same family. The PVE of each individual locus ranged from 7.92% (qMIC-5-1) to 22.95% (qMIC-25-1), with an average of 14.06%. For 15 of the 27 QTLs, G. mustelinum alleles reduced MIC and therefore contributed to finer fiber (Table 2). Eighteen of the 27 QTLs could also be identified by the MCIM method of QTLNetwork (Table 2).

Consistency of QTLs across Families

A total of 211 of the 218 SSR marker loci were segregating in two or more families; consequently, two-way ANOVA was performed to identify marker-trait associations, so as to evaluate their consistency among different families. Significant (P < 0.001) among-family G effects were identified at 13 and 10 loci for STR and MIC, respectively (Table S2). For STR, the 13 loci appear to represent eight non-overlapping genomic regions, and QTLs were identified in four regions in within-family analysis (qSTR-5-1, qSTR-6-1, qSTR-19-2, and qSTR-23-1). For MIC, the 10 loci represent three non-overlapping genomic regions, for which QTLs were detected in three regions in within-family analysis (qMIC-5-1, qMIC-7-1, and qMIC-15-1, Table 2, Table S2).

For MIC, a total of five loci were significant (P < 0.001) for G × F interactions (Table S3). QTLs were detected at all five loci (qMIC-4-1, qMIC-4-2, qMIC-10-1, qMIC-15-2, qMIC-19-3, and qMIC-25-2; Table 2, Table S3) in different segregating families. No locus was significant for G × F interactions for STR.

Epistatic QTLs and Their Interactions with Environments

A total of 13 epistatic QTLs were identified with significant additive × additive (AA) effects (P < 0.001), most of which (10/13) involved loci that were not linked to any main-effect QTLs (Table 3). For STR, the interaction between a region on Chr5 and another region on Chr19 was identified in both BC3F2.3 and joint analysis in POP34 simultaneously. G. mustelinum alleles increased STR (with negative AA effects) for three epistatic QTLs, and alleles from G. mustelinum decreased MIC (with positive AA effects) for five epistatic QTLs (Table 3). The interactions between epistatic QTLs with environment were listed in Table 3.

Discussion

Cultivated Gossypium hirsutum has a narrow gene pool, having experienced genetic bottlenecks during polyploid formation and divergence from its sister polyploid species, as well as during...
FIGURE 2 | QTLs of fiber strength (STR) and fineness (Micronaire, MIC). The solid black bars indicated QTLs of STR, and the green bars filled with slashes indicated QTLs of MIC.
| QTL<sup>a</sup>  | Generation | Locus     | Family | $R^2$ (%)<sup>b</sup> | $A^b$ | $D^b$ | D/A ratio<sup>b</sup> | Mode of action<sup>c</sup> |
|-----------------|------------|-----------|--------|----------------------|-------|-------|----------------------|-----------------------------|
| qSTR-1-1<sup>+</sup> | BC<sub>0F2;3</sub> | CIR009     | POP20  | 12.56                | −1.46 | −0.80 | 0.55                | A                           |
| qSTR-2-1<sup>+</sup> | BC<sub>0F2;3</sub> | BNL3971    | POP20  | 12.17                | −1.06 | −0.08 | 0.08                | A                           |
| qSTR-5-1<sup>+</sup> | BC<sub>0F2;3</sub> | DPL0674    | POP20  | 11.14                | −1.24 | −0.43 | 0.34                | A                           |
| qSTR-6-1         | BC<sub>0F2;3</sub> | DPL0241    | POP04  | 14.30                | −0.13 | −1.54 | 11.51              | H                           |
| qSTR-7-1<sup>+</sup> | BC<sub>0F2;3</sub> | NAU5433    | POP32  | 16.69                | 1.90  | 0.71  | 0.38                | A                           |
| qSTR-11-1<sup>+</sup> | BC<sub>0F2;3</sub> | BNL3442    | POP16  | 14.04                | 1.23  | 0.28  | 0.23                | A                           |
| qSTR-12-1        | BC<sub>0F2;3</sub> | NAU3377b   | POP16  | 18.12                | 1.66  | –     | –                  | –                           |
| qSTR-19-1<sup>+</sup> | BC<sub>0F2;3</sub> | JESP0800   | POP03  | 16.32                | 1.47  | 0.86  | 0.59                | A                           |
| qSTR-19-2<sup>+</sup> | BC<sub>0F2;3</sub> | BNL3811    | POP15  | 8.96                 | −0.84 | 0.00  | 0.00                | A                           |
| qSTR-22-1        | BC<sub>0F2;3</sub> | BNL3811    | POP15  | 11.70                | −0.89 | −0.41 | 0.47                | A                           |
| qSTR-22-2<sup>+</sup> | BC<sub>0F2;3</sub> | DPL0417    | POP01  | 18.34                | −1.12 | −0.13 | 0.11                | A                           |
| qSTR-23-1<sup>+</sup> | BC<sub>0F2;3</sub> | CIR048     | POP10  | 10.28                | −0.62 | −0.08 | 0.12                | A                           |
| qSTR-25-1<sup>+</sup> | BC<sub>0F2;3</sub> | BNL3264    | POP17  | 14.53                | −0.68 | 0.57  | −0.84               | A                           |
| qMIC-1-1         | BC<sub>0F3;4</sub> | TMH010    | POP10  | 16.00                | 0.00  | 2.08  | 1.04                | H                           |
| qMIC-4-1         | BC<sub>0F3;3</sub> | DPL01987a  | POP32  | 15.58                | 0.25  | −0.37 | −0.87               | D                           |
| qMIC-4-2         | BC<sub>0F3;3</sub> | MUSB1050   | POP32  | 21.45                | 0.41  | −0.08 | −0.20               | A                           |
| qMIC-4-3         | BC<sub>0F3;3</sub> | DPL0667    | POP35  | 12.26                | −0.05 | −0.31 | 0.51                | H                           |
| qMIC-5-1         | BC<sub>0F2;2</sub> | BNL2656    | POP15  | 8.92                 | −0.20 | −0.02 | 0.10                | A                           |
| qMIC-5-2         | BC<sub>0F2;2</sub> | BNL2656    | POP15  | 14.09                | −0.21 | −0.06 | 0.30                | A                           |
| qMIC-7-1         | BC<sub>0F2;2</sub> | BNL3498    | POP15  | 8.86                 | −0.21 | −0.12 | 0.56                | A                           |
| qMIC-7-2         | BC<sub>0F2;2</sub> | MUCA616    | POP32  | 15.91                | 0.63  | 0.26  | 0.41                | A                           |

(Continued)
domestication, dispersal by humans, and scientific breeding. The domestication of a comparatively small subgroup of wild species and, in more recent years, over-exploitation of only a few genetic backgrounds in breeding programs of modern Upland cotton by crossing a few closely-related genotypes repeatedly to develop new cultivars has resulted in a genetically depauperate cotton germplasm. Slow genetic progress in improving fiber productivity and quality is indicative that many favorable alleles have reached fixation in the elite gene pool. The deficiency of genetic variation in current germplasm has enhanced the difficulty for breeders to provide low-cost intrinsic genetic solutions to fiber quality traits such as resistance to abiotic and biotic hazards or new needs in fiber quality or agronomic traits.

Because the narrowing of the cotton genetic base, new sources of genetic variation need to be introduced into the cotton gene pool to ensure future success in breeding new cotton cultivars. As a wild tetraploid cotton species diverged far from G. hirsutum with cultivated cotton, it may harbor elite alleles beneficial to the improvement of fiber quality traits in Upland cotton (Wendel and Cronn, 2003) but sexually compatible with cultivated cotton, G. mustelinum may harbor elite alleles beneficial to the improvement of fiber quality traits in Upland cotton (Alves et al., 2013).

In this research, advanced backcross QTL (AB-QTL) analysis was carried out based on SSR markers and phenotypic data.

Table 2: Continued

| QTL*  | Generation | Locus | Family | R^2 (%)b | A^b  | D^b  | D/A ratio^b | Mode of action^c |
|-------|------------|-------|--------|----------|------|------|-------------|-----------------|
| qMIC-8-1 | BCf2 | CIR343 | POP11 | 16.91 | 0.59 | 0.38 | 0.65 | A |
| qMIC-9-1 | BCf2 | BNL3985 | POP11 | 15.10 | 0.59 | – | – | – |
| qMIC-10-1 | BCf2 | DPL0385 | POP11 | 14.54 | 0.57 | – | – | – |
| qMIC-10-2 | BCf2 | BNL2631 | POP35 | 13.00 | –0.14 | 0.18 | –1.27 | D |
| qMIC-12-1 | BCf2 | DPL0866 | POP11 | 17.45 | 0.56 | 0.43 | 0.77 | A |
| qMIC-13-1 | BCf2 | BNL1438 | POP32 | 16.55 | 0.48 | –0.17 | –0.36 | A |
| qMIC-15-1 | BCf2 | BNL2646 | POP15 | 10.81 | 0.16 | – | –0.09 | –0.56 | A |
| qMIC-15-2 | BCf2 | NAU4045 | POP27 | 13.49 | –0.30 | – | – | – |
| qMIC-16-1 | BCf2 | DPL0385b | POP27 | 10.17 | 0.30 | 0.12 | 0.35 | A |
| qMIC-16-2 | BCf2 | NAU5024 | POP15 | 9.12 | –0.22 | – | –0.04 | 0.18 | A |
| qMIC-19-1 | BCf2 | NAU3205 | POP09 | 17.13 | 0.11 | –0.42 | –4.04 | H |
| qMIC-19-2 | BCf2 | BNL5355a | POP09 | 13.29 | 0.17 | – | –0.29 | –1.70 | D |
| qMIC-19-3 | BCf2 | CIR212 | POP16 | 13.15 | –0.18 | –0.20 | 1.13 | D |
| qMIC-24-1 | BCf2 | NAU3605 | POP02 | 11.15 | 0.40 | 0.27 | 0.68 | A |
| qMIC-25-1 | BCf2 | JESP1224 | POP01 | 18.73 | 0.16 | –0.41 | –2.64 | D |
| qMIC-25-2 | BCf2 | BNL4001b | POP01 | 19.28 | 0.83 | 0.46 | 0.55 | A |
| qMIC-25-3 | BCf2 | BNL4001b | POP17 | 22.95 | 0.13 | –0.53 | –3.94 | H |
| qMIC-25-4 | BCf2 | BNL3264 | POP17 | 20.30 | 0.14 | –0.46 | –3.24 | H |
| qMIC-26-1 | BCf2 | STS511 | POP11 | 10.49 | 0.27 | – | – | – |
| qMIC-26-2 | BCf2 | BNL4001b | POP17 | 22.57 | –0.01 | –0.41 | 42.30 | H |
| qMIC-26-3 | BCf2 | BNL1163 | POP10 | 15.08 | 0.53 | 0.69 | 1.30 | D |
| qMIC-26-4 | BCf2 | BNL2725 | POP15 | 8.79 | –0.16 | –0.05 | 0.33 | A |
| qMIC-26-5 | BCf2 | BNL341 | POP12 | 11.35 | –0.19 | –0.07 | 0.34 | A |
| qMIC-26-6 | BCf2 | BNL341 | POP15 | 9.20 | –0.15 | 0.01 | –0.05 | A |

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

* indicating that the QTL was also detected by QTLNetwork.

1Quantitative 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.

2Modes of gene action are indicated by: A, additivity; D, dominance; H, overdominance. Missing values correspond to dominant marker loci.
collected from three generations of BC$_3$-derived families with introgression from G. mustelinum. Phenotypic assessment of the advanced backcross populations indicated significant segregation for STR and MIC, which indicated that both negative and positive alleles existed for each trait in both parents (Table 1, Figure 1).

Since the wild parent, G. mustelinum cannot produce spinnable fiber, it is not a surprise to find out that for STR, many BC$_3$-derived families (16 of 21 in BC$_3$F$_2$, five of 12 in BC$_3$F$_3$, and three of 12 in BC$_3$F$_4$) showed mean STR values lower than the recurrent G. hirsutum parent, PD94042; for MIC, there were five BC$_3$-derived families in each generation having higher mean MIC values or coarser fiber than the recurrent parent. This “negative” transgression, yielding a poorer phenotype than that of the recurrent parent, suggests that interspecific hybridization formed some desirable new gene combinations, which was also found in previous reports on introgression of G. barbadense (Chee et al., 2005a,b; Dray et al., 2005), G. tomentosum (Zhang et al., 2011), and G. darwinii (Wang et al., 2012). As suggested by previous research (Jiang et al., 2000), gene transfer between gene pools is a significant consequence of interspecific hybridization, which will increase selectable genetic variation and introduce genes for adaptive traits. These results support the hypothesis that fiber quality improvement for Upland cotton may be accomplished by introgressing elite genes from G. mustelinum and other tetraploid cotton species.

A total of 15 and 27 non-overlapping QTLs were mapped in one-way analysis of variance with PVE of 13.86 and 14.06% on average for STR and MIC respectively. The effects of many QTLs showed good reproducibility, with 18 of the 42 QTLs detected at least twice near the same markers in different generations/families or near linked markers in the same family. In addition, 28 of the 42 QTLs were also identified by QTLNetwork (Table 2). The detection of QTLs near different markers at corresponding chromosomal locations or across various generations/families with different methods further supports the likelihood that these QTLs are real.

Alleles from the wild G. mustelinum increased STR for 53% (eight) of 15 QTLs and decreased MIC (conferred finer fiber) for 56% (15) of 27 QTLs. These QTLs are of great importance to be deeply exploited to transfer elite genes from G. mustelinum into Upland cotton. Efforts to improve STR and MIC are now in progress by constructing near-isogenic lines allowing these alleles

---

**Table 3** Estimated epistasis and epistasis × environment interaction effects of QTLs for fiber strength and fineness.

| Trait | Gen./Env. | Family | QTL$^b$ | Chromosome | Interval | QTL$^b$ | Chromosome | Interval | AA$^c$ | P-Value | $h^2$(aa) | $h^2$(aae) |
|-------|-----------|--------|---------|------------|----------|---------|------------|----------|-------|---------|----------|-----------|
| STR BC$_3$F$_2$ | POP20 | qSTR-1-1 | Chr1 | MUS166-CIR009 | qSTR-2-1 | Chr2 | BNL3971-DPL0674 | 1.52 | 0.000000 | 9.79 |
| Joint POP20 | qSTR-1-1 | Chr5 | DPL0241-BNL3029 | – | Chr19 | CIR212-DPL0444 | 1.59 | 0.000151 | 11.68 |
| Joint POP31 | – | Chr3 | BNL3441-BNL3267 | – | Chr19 | BNL3715-CIR212 | –1.24 | 0.000013 | 9.99 | 1.43 |
| Joint POP20 | – | Chr16 | DPL0501-NAU2862 | – | Chr18 | BNL193-BNL243 | –0.98 | 0.000000 | 5.50 | 0.44 |
| Joint POP11 | – | Chr20 | BNL169-ST5324 | – | Chr26 | NAU3862-NAU1119 | 0.99 | 0.000000 | 10.25 | 0.13 |
| Joint POP17 | qSTR-25-1 | Chr25 | BNL3264-BNL4001b | – | Chr24 | NAU3865-DPL0624 | –0.93 | 0.000000 | 3.71 | 0.71 |

**Notes:***Joint: results obtained based on combined data of the BC$_3$F$_2$, BC$_3$F$_3$, and BC$_3$F$_4$ generations.**

**Table 4** Estimated joint effects of QTLs in BC$_3$F$_2$-derived families.

| Joint POP1 | Chr1 | BNL1350-BNL2646 | Chr5 | CIR102-DPL0241 | 0.41 | 0.000006 | 3.68 |
| Joint POP2 | – | BNL1606-NAU3639 | – | CIR102-DPL0241 | 0.63 | 0.000001 | 21.05 |
| Joint POP7 | – | BNL1434-BNL3972 | qMIC-4-3 | CIR102-DPL0241 | –0.20 | 0.000001 | 6.41 | 2.71 |
| Joint POP6 | – | BNL3400-CIR102 | Chr4 | DPL0085-DPL0676 | –0.20 | 0.000001 | 6.41 | 2.71 |
| Joint POP2 | – | NAU4900-CIR354b | – | CIR102-DPL0241 | 0.20 | 0.000001 | 4.13 | 2.13 |
| Joint POP1 | – | NAU3639-NAU5443 | – | CIR102-DPL0241 | 0.20 | 0.000001 | 4.13 | 2.13 |
| Joint POP1 | – | BNL3511-DPL0378 | – | CIR102-DPL0241 | –0.27 | 0.000001 | 3.72 | 2.18 |
| Joint POP2 | – | BNL3903-BNL3811 | – | CIR102-DPL0241 | 0.27 | 0.000000 | 5.85 | 1.32 |
| Joint POP2 | – | BNL2589-NAU39074 | – | CIR102-DPL0241 | 0.29 | 0.000000 | 1.15 | 6.42 |
from *G. mustelinum* to be more easily manageable in cotton breeding programs.

In addition to main-effect QTLs, epistatic QTLs were identified for STR and MIC using data collected over three different generations (Tables 3). The results indicated that both epistatic QTLs and main-effect play key genetic roles in STR and MIC (Tables 2, 3). Epistatic QTLs are more complicated to manage compared to utilization of main-effect alleles, naturally at least twice the difficulty will be encountered in testing and introgressing of epistatic QTLs. Furthermore, it is interesting that the majority of epistatic interactions (10/13) were identified between genetic loci that were not linked to any QTL (Table 3), similar to our previous results derived from fiber length, where 14 of 17 epistatic QTLs involved loci not linked to any main-effect QTLs (Wang et al., 2017), suggesting high complexity of fiber quality inheritance. In joint analysis, effects of epistasis × environment were detected for both STR and MIC (Table 3). Although, generally the PVE of epistasis × environment was smaller than that of epistasis, epistasis × environment interactions may add to the difficulty of breeding.

Since more than one family was often segregating for the same chromosomal segment, it was possible to explore genetic background effects on introgressed chromatin. For STR and MIC, significant (*P* < 0.001) among-family *G* effects were identified at 13 and 10 loci (Table S2), with four and three loci revealing QTLs in within-family analysis for STR (*qSTR-5-1, qSTR-6-1, qSTR-19-2, and qSTR-23-1*) and MIC (*qMIC-5-1, qMIC-7-1, and qMIC-16-1*). Some among-family *G* effects demonstrated good reproducibility, with three of 13 for STR, and two of 10 for MIC detected in different generations (Table S2). A total of five loci were significant (*P* < 0.001) among-family *F* interactions for MIC (Table S3), with QTLs detected in all the five loci (*qMIC-4-1, qMIC-4-2, qMIC-10-1, qMIC-15-2, qMIC-19-3, and qMIC-25-2*; Table 2, Table S3). The most extreme case of *G* × *F* interaction was detected at the locus MUSB1050 on Chr4 (Table 2, Table S3). In family POP32, alleles from *G. mustelinum* at this locus conferred a decrease of 0.41 for additive effect in MIC that accounted for 21.45% of PVE. Interestingly, this same locus conferred an additive increase of 0.09 for MIC that accounted for 10.78% of PVE in family POP35; this locus also segregated in family POP11, POP10, and POP27 but showed no significant association with MIC.

This study also adds to prior information on the significant influence of the tetraploid D-subgenome on fiber quality traits, although the D-subgenome was derived from a diploid ancestor without the ability to produce spinnable fiber (Jiang et al., 1998; Chee et al., 2005b; Zhang et al., 2008, 2011). In this current study, among the 42 QTLs affecting STR and MIC traits, the D-subgenome (20) had slightly fewer QTLs than the A-subgenome (22). Considering the QTLs for fiber elongation and fiber length detected in our previous reports (Wang et al., 2016a, 2017), 73 QTLs were identified in the D-subgenome, more than the A-subgenome (58), which collectively supports the finding that for fiber quality traits, more QTLs occurred on the D-subgenome than the A-subgenome (Jiang et al., 1998; Paterson et al., 2003).

With this fourth report on QTLs for fiber traits from *G. mustelinum*, this series of papers collectively describes six fiber quality traits investigated in 21 BC1F2, 12 BC1F2, 3, and 12 BC1F2, 4 families, namely fiber elongation (EL), fiber uniformity index (UI), upper-half mean length (UHM), short fiber content (SFC), STR, and MIC. Mean values of some families outperformed the recurrent *G. hirsutum* parent, PD94042, for each trait in each generation; likewise, many individual plants/lines showed superior fiber quality performance than the *G. hirsutum* parent, showing promise that our goal of introgressing alleles from *G. mustelinum* to improve Upland cotton may work.

Two loci (JESPR224 and BNL3264) may be of special interest in this series of reports. Two QTLs were detected near the locus JESPR224 (*qMIC-25-1* in POP01 in BC1F2, and *qSTR-25-2* in POP02 in BC1F2), where *G. mustelinum* alleles decreased MIC and increased STR. Four QTLs (*qUHM-25-1, qSTR-25-1, qMIC-25-1, and EL25.1*) were detected near the loci BNL3264 on Chr21 in POP17, for which *G. mustelinum* alleles increased EL, UHM, and STR and decreased MIC (Wang et al., 2016a, 2017). These QTLs hold promise for improving cotton fiber elongation, length, strength and fineness simultaneously through marker-assisted selection.

More co-localizations of QTLs with both desirable and undesirable effects on different traits were observed. For instance, two QTLs (*qELO-19-1* and *qSTR-19-2*) were detected in POP15 near the locus BNL3811 on Chr19, for which *G. mustelinum* alleles increased STR but decreased EL; three QTLs (*qMIC-12-1, qSFC-12-2, and qUI-12-3*) were detected near the locus DPL0866 on Chr12 in POP11, for which *G. mustelinum* alleles decreased MIC but increased SFC and decreased UI; four QTLs (*qUI-10-1, qSFC-10-2, qMIC-10-1, and qELO-10-1*) were detected near the locus JESPR6 on Chr10, for which *G. mustelinum* alleles increased EL but they also increased MIC and SFC and decreased UI (Wang et al., 2016a, 2017). The co-location of QTLs with opposite effects on different traits indicates the difficulty of improving diverse traits simultaneously in breeding programs, and may also account for the challenges that have been faced with using exotic germplasm in the absence of DNA marker information.

Building on prior work on *G. barbadense*, *G. tomentosum*, and *G. darwinii*, QTL mapping involving introgression of *G. mustelinum* alleles offers new allelic variation to the Upland cotton gene pool; in addition, the new germplasm created here offers an opportunity for the cotton community to explore *G. mustelinum* alleles in an elite cultivated background, and also provided materials potentially useful in cotton breeding programs. Wild cotton species represent a repository of divergent and in some cases favorable alleles for a variety of traits including fiber quality. The sixth and seventh tetraploid cotton species found and confirmed, *G. ekmanianum* Wittmack (endemic to the Dominican Republic; Krapovickas and Seijo, 2008; Wendel and Grover, 2015) and *Gossypium* sp. nov. (found from two islands, Wake and Peale in the Wake Atoll in the Pacific Ocean; Wendel and Grover, 2015) offer additional scope for exploration of gene introgression.
AUTHOR CONTRIBUTIONS
AP conceived and designed the experiments; PC and OM oversaw crossing, phenotyping, and genotyping, which were performed by BW, ZhiZ, ZheZ and EL; DJ oversaw fiber analysis; BW, XD, LS, and TS analyzed the data; BW, AP, and PC wrote the paper.

FUNDING
This work was supported by National Science Foundation (PFI award IIP 0917856; AIR award IIP 1127755), Cotton Incorporated, Key Research and Development Project of Jiangsu Province, China (Modern Agriculture, BE2015353), the Applying Basic Research Fund of Nantong City-Agricultural and Social Development (MS12016043).

ACKNOWLEDGMENTS
We thank Jennifer McCurdy for help in field work and Cornelia Lemke for technical assistance in the lab.

SUPPLEMENTARY MATERIAL
The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2017.01848/full#supplementary-material

REFERENCES
Alves, M., Barroso, P., Ciampi, A., Hoffmann, L., Azevedo, V., and Cavalcante, U. (2013). Diversity and genetic structure among subpopulations of Gossypium mustelinum (Malvaceae). Genet. Mol. Res. 12, 597–609. doi: 10.4238/2013.February.27.9
Chee, P., Draye, X., Jiang, C. X., Decanini, L., Delmonte, T. A., Bredhauer, R., et al. (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. doi: 10.1007/s00122-005-0263-z
Chee, P. W., Draye, X., Jiang, C. X., Decanini, L., Delmonte, T. A., Bredhauer, R., et al. (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. doi: 10.1007/s00122-005-2062-0
Draye, X., Chee, P., Jiang, C. X., Decanini, L., Delmonte, T. A., Bredhauer, R., et al. (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. doi: 10.1007/s00122-005-2061-1
Estur, G., and Knappe, M. (2007). Cotton Exporter’s Guide. Geneva: The International Trade Centre (ITC), Palais des Nations.
Jiang, C. X., Chee, P. W., Draye, X., Morrell, P. L., Smith, C. W., and Paterson, A. H. (2000). Multilocus interactions restrict gene introgression in interspecific populations of polyploid Gossypium (cotton). Evolution 54, 798–814. doi: 10.1111/j.0014-3820.2000.tb00811.x
Jiang, C. X., Wright, R. J., El-Zik, K. M., and Paterson, A. H. (1998). Polyploid formation created unique avenues for response to selection in Gossypium (cotton). Proc. Natl. Acad. Sci. U.S.A. 95, 4419–4424. doi: 10.1073/pnas.95.8.4419
Krapivovska, A., and Seijo, G. (2008). Gossypium ekmanianum (Malvaceae), algodon silvestre de la Republica Dominicana. Bonplandia 17, 55–63. Available online at: http://ibone.unne.edu.ar/objetos/up/documentos/bonplandia/public/17/1_155_63.pdf
McCouch, S., Cho, Y., Yano, M., Paul, E., Blinestub, M., Morishima, H., et al. (1997). Report on QTL nomenclature. Rice Genet. NewsL. 14:11.
Montalvo, J. R. (2005). Relationships between micronaire, fineness, and maturity. I. Fundamentals. J. Cotton Sci. 9, 81–88. Available online at: http://www. leadershipcotton.org/journal/2005-09-02/upload/jc09-081.pdf
Naylor, G. R., Delhom, C. D., Cui, X., Gourot, J.-P., and Rodgers, J. (2014). Understanding the influence of fiber length on the High Volume Instrument™ measurement of cotton fiber strength. Textile Res. J. 84, 979–988. doi: 10.1177/0000652115531318
Paterson, A. H., Deverna, J. W., Lanini, B., and Tanksley, S. D. (1990). Fine mapping of quantitative trait loci using selected overapping recombinant chromosomes, in an interspecies cross of tomato. Genetics 124, 735–742.
Paterson, A. H., Saranga, Y., Mentz, M., Jiang, C. X., and Wright, R. J. (2003). QTL analysis of genotype x environment interactions affecting cotton fiber quality. Theor. Appl. Genet. 106, 384–396. doi: 10.1007/s00122-002-1025-y
SAS Institute (1999). SAS/STAT User’s Guide, Version 8. Cary, NC: SAS Institute.
Self, S. G., and Liang, K.-Y. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. J. Am. Stat. Assoc. 82, 605–610. doi: 10.1080/01621459.1987.10478472
Wang, B., Draye, X., Zhang, Z., Zhuang, Z., May, O. L., Paterson, A. H., et al. (2016a). Advanced backcross QTL analysis of fiber elongation in a cross between Gossypium hirsutum and G. mustelinum. Crop Sci. 56, 1760–1768. doi: 10.2135/cropsci2015.12.0753
Wang, B., Draye, X., Zhuang, Z., Zhang, Z., Liu, M., Lubbers, E. L., et al. (2017). QTL analysis of cotton fiber length in advanced backcross populations derived from a cross between Gossypium hirsutum and G. mustelinum. Theor. Appl. Genet. 6, 1297–1308. doi: 10.1007/s00122-017-2889-1
Wang, B., Liu, L., Zhang, D., Zhuang, Z., Guo, H., Qiao, X., et al. (2016b). A genetic map between Gossypium hirsutum and the Brazilian endemic G. mustelinum and its application to QTL mapping. G 6, 1673–1685. doi: 10.1534/g3.116.029116
Wang, B., Nie, Y. C., Lin, Z. X., Zhang, X. L., Liu, J. J., and 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. doi: 10.1007/s00122-012-1911-x
Wendel, J. F., and Cronn, R. C. (2003). Polyploidy and the evolutionary history of cotton. Adv. Agron. 78, 139–186. doi: 10.1016/S0065-2113(02)78004-8
Wendel, J. F., and Grover, C. E. (2015). “Taxonomy and evolution of the genus Gossypium,” in Cotton, 2nd Edn, eds D. Fang and R. Percy (Madison, WI: American Society of Agronomy, Inc.; Crop Science Society of America, Inc.; and Soil Science Society of America, Inc.).
Yung, J. H., Hu, C. C., Hu, H., Yu, R. D., Xia, Z., Ye, X. Z., et al. (2008). QTLNetwork: mapping and visualizing genetic architecture of complex traits in experimental populations. Bioinformatics 24, 721–723. doi: 10.1093/bioinformatics/btm494
Zhang, H. B., Li, Y., Wang, B., and Chee, P. W. (2008). Recent advances in cotton genomics. Intern. J. Plant Genomics 2008:742304. doi: 10.1155/2008/742304
Zhang, Z. S., Rong, J. K., Waghmare, V. N., Chee, P. W., May, O. L., Wright, R. J., et al. (2011). QTL alleles for improved fiber quality from a wild Hawaiian cotton, Gossypium tomentosum. Theor. Appl. Genet. 123, 1075–1088. doi: 10.1007/s00122-011-1649-x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Wang, Zhaung, Zhang, Draye, Shuang, Shehzad, Lubbers, Jones, May, Paterson and Chee. This is an open-access article distributed under the terms of the Creative Commons Attribution License (C C BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.