Theor Appl Genet (2005) 111: 772–781
DOI 10.1007/s00122-005-2062-0

ORIGINAL PAPER

**Peng W. Chee · Xavier Draye · Chun-Xiao Jiang
Laura Decanini · Terrie A. Delmonte
Robert Bredhauer · C. Wayne Smith
Andrew H. Paterson**

# Molecular dissection of phenotypic variation between *Gossypium hirsutum* and *Gossypium barbadense* (cotton) by a backcross-self approach: III. Fiber length

Received: 9 November 2004 / Accepted: 2 May 2005 / Published online: 28 June 2005


**Abstract** A backcross-self population from a cross between *Gossypium hirsutum* and *G. barbadense* was used to dissect the molecular basis of genetic variation governing 15 parameters that reflect fiber length. Applying a detailed restriction fragment length polymorphism (RFLP) map to 3,662 $BC_3F_2$ plants from 24 independently derived $BC_3$ families, we detected 28, nine, and eight quantitative trait loci (QTLs) for fiber length, length uniformity, and short fiber content, respectively. For eight, six, and two chromosomal regions containing quantitative trait loci (QTLs) for fiber length, length uniformity, and short fiber content (respectively), two-way analysis of variance showed a significant ($P < 0.001$) among-family genotypic effect. A total of 13, two, and four loci showed genotype × family interaction, illustrating some of the complexities that are likely to be faced in introgression of exotic germplasm into the gene pool of cultivated cotton. Co-location of many QTLs for fiber length, length uniformity, and short fiber content accounted for correlations among these traits, while the discovery of many QTLs unique to each trait suggests that maximum genetic gain will require breeding efforts that target each trait (or an index including all three). The availability of DNA markers linked to *G. barbadense* QTLs identified in this and other studies promise to assist breeders in transferring and maintaining valuable traits from exotic sources during cultivar development.

**Electronic Supplementary Material** Supplementary material is available for this article at http://dx.doi.org/10.1007/s00122-005-2062-0

Communicated by F. Salamini

P. W. Chee · A. H. Paterson
Plant Genome Mapping Laboratory,
University of Georgia, Athens, Georgia 30602, USA

P. W. Chee · A. H. Paterson
Department Crop and Soil Sciences,
University of Georgia, Tifton, Georgia 31793, USA

P. W. Chee · X. Draye · C.-X. Jiang · L. Decanini
T. A. Delmonte · R. Bredhauer · C. W. Smith
A. H. Paterson (✉)
Department Soil and Crop Sciences, Texas A&M University,
College Station, Texas, 77843, USA
E-mail: paterson@uga.edu
Tel.: +706-583-0162
Fax: +706-586-0160

*Present address*: X. Draye
Université catholique de Louvain,
Place Croix du Sud 2/11, 1348 Louvain-la-Neuve, Belgium

## Introduction

Cotton fiber quality is defined by physical properties that relate to its spinnability into yarn, contribute to textile performance, and enhance end-product quality. One of the most important aspects of fiber quality is length, including average length of fibers, length uniformity of the population of fibers, and content of short fibers. Longer fibers can be processed at greater efficiencies and produce finer (and therefore higher quality) yarns, while shorter fibers require increased twisting during spinning, causing low-strength, poor-quality yarns (Perkins et al. 1984). Further, low length uniformity and high short fiber content are associated with increased manufacturing waste and decreased spinning efficiency during yarn processing. As a result of the demand for improved fiber quality, the US cotton marketing system imposes premiums and discounts for fiber length and other related qualities (Jost 2002). Poor fiber length can result in severe price discounts for cotton growers, as has been experienced in recent years by growers in Texas and Georgia, two of the largest cotton-producing states in the USA.

Cotton fiber production is dominated by two tetraploid species that are each cultivated for somewhat different characteristics. *Gossypium hirsutum*, commonly known as Upland cotton, is the most widely grown species because of its high-yield potential and adaptation to diverse environmental conditions and production systems. On the other hand, *G. barbadense*, also known as Pima, Egyptian, Sea Island, and Extra Long Staple cotton, has a fiber quality much superior to that of Upland cotton. However, the narrow range of adaptation of this latter species—it is limited to irrigated regions in arid zones of Arizona and California—precludes its use in most of the US "Cotton Belt". Nevertheless, the unique fiber properties of *G. barbadense* make it an ideal candidate for providing new genetic variation useful for improving fiber quality in Upland cotton.

In a companion paper, we describe the development of an advanced-generation interspecific backcross population by utilizing a cultivated *G. barbadense* as the donor parent to derive 24 independent $BC_3F_2$ families, with each segregating for a unique set of *G. barbadense* chromosome segments in an otherwise *G. hirsutum* genetic background (Chee et al. 2005). We also describe the use of DNA markers to characterize introgressed chromatin segregating in each family and identify regions associated with quantitative trait loci (QTLs) affecting fiber elongation and fiber fineness. This paper reports the results of QTL analysis for fiber length-related properties in the backcross population. In addition to identifying the subgenomic distribution of genetic loci responsible for fiber length traits, we also examined the inter-relationship between QTLs associated with the three most important fiber length parameters, average fiber length, length uniformity, and short fiber content.

## Materials and methods

### Population development and field evaluation

A backcross-self-mating design was used in this study. About 100 *Gossypium hirsutum* cv. Tamcot 2111 plants were pollinated by *G. barbadense* cv. Pima S6 to produce $F_1$ hybrid plants. One interspecific hybrid plant from each cross was then backcrossed three times to Tamcot 2111 to yield $BC_3F_1$s. A total of 24 lineages led to the production of at least one $BC_3F_1$ plant, which was self-pollinated to generate 24 $BC_3F_2$ families ranging in size from 33 to 191 plants per family, for a grand total of 3,662 $BC_3F_2$ progeny. The $BC_3F_2$ progenies were space-planted together with ten entries of each parent as checks in a field near College Station, Texas, under fertilization and irrigation regimes consistent with commercial cotton production (for more detailed descriptions, see Chee et al. 2005). Although all progeny were grown to maturity and DNA sampled, a number of plants were lost due to sterility or otherwise produced insufficient lint for fiber analysis. Fiber samples were harvested only from 2,976 progenies, ranging in size from 22 to 172 plants per family. Seedcotton from all bolls of an individual $BC_3F_2$ plant were hand-harvested and ginned on a saw gin.

Fifteen traits reflecting fiber length parameters were determined by the Cotton Incorporated Textile Services Laboratory (Cotton Incorporated, Cary, N.C.). Eight traits were measured using the High-Volume Precision Instrument (HVI), including upper-half mean length (HVuhm), staple length (HVstapl), 2.5% span length (HVsl2.5), 50% span length (HVsl50), percentage short fiber by weight (HVsfc), uniformity index (HVui), uniformity ratio (HVur), and coefficient of variation (CV) of length by number (HVlnCV). Seven traits were measured by the Advanced Fiber Information System (AFIS) instrument, including length by weight (Lw), length by number (Ln), upper quartile length by weight (UQLw), short fiber content by weight (SFCw), short fiber content by number (SFCn), CV of length by number (LnCV), and CV of length by weight (LwCV). For more detailed descriptions of the measurement methods for each trait, see Steadman (1997).

### Genotyping and data analysis

Laboratory techniques were as described in Jiang et al. (2000a). The genome composition of the $BC_3F_1$ plants was inferred based on genotyping the pooled DNA of 20–30 $BC_3F_2$ plants with 262 restriction fragment length polymorphism (RFLP) markers selected for even representation of the cotton genome map (Reinisch et al. 1994; Jiang et al. 1998). The markers that detected introgression in the $BC_3F_1$ were then used to evaluate the entire $BC_3F_2$ family for which they were segregating. A subset of 127 RFLP markers proved sufficient to monitor all the segregating regions found in the $BC_3F_2$ families. Regions containing *G. barbadense* introgression that were segregating in each $BC_3F_2$ family were monitored using graphical genotypes (Young and Tanksley 1989) constructed based on a detailed genetic map comprised of over 2,500 loci (Rong et al. 2004).

One-way ANOVA (SAS, ver. 8, GLM; SAS Institute 1999), the modes of gene action (*a* and *d*), and the proportion of phenotypic variance explained by a QTL were calculated for every marker locus segregating within each $BC_3F_2$ family as described in Chee et al. (2005). For loci that were segregating in two or more families, two-way mixed model variance analyses were also performed, using the MIXED procedure of the SAS ver. 8 package. The variance analysis model included genotype (G) as a fixed factor and family (F) and genotype × family (G×F) interaction as random factors. Model parameters were estimated using the residual maximum likelihood (REML) method. The marker-trait association (genotype factor) was tested with an *F* statistic, using a general Satterthwaite approximation for the denominator degrees of freedom (SAS Institute

1999). A likelihood-ratio statistic (CHISQ) was performed for the G×F interaction (Self and Liang 1987). G effects and G×F interactions were considered to be significant if $P < 0.001$.

## Results

### Measures and correlations

The correlations among the 15 fiber length-related parameters are shown in Table 1. With such a large sample size (varied from 2,979 to 2,993 individuals due to missing data), virtually all correlations were statistically significant. We were especially interested in identifying traits for which QTL data were nearly identical, so as to only present non-redundant information. Traits that correlated at absolute values of 0.93 or greater (Table 1) showed nearly identical sets of QTLs. We present full data for the eight most informative phenotypes representative of each such closely correlated group to reflect the three basic fiber length properties, which we will refer to as average fiber length (Lw, HVuhm, and HVsl2.5), fiber length uniformity (HVui, HVur, and HVlnCV), and short fiber content (HVsfc and SFCn). Among the eight phenotypes, the *G. barbadense* parent confers a more desirable phenotype for most traits, the exceptions being two measures of fiber length uniformity, HVlnCV and HVur.

The phenotypic distributions of the $BC_3F_2$ populations, together with the mean values of each $BC_3$ family and the parents, are presented in Fig. 1. Transgressive segregation for all phenotypes was evident in the $BC_3F_2$ progenies. However, the mean for each $BC_3F_2$ family varied markedly for each phenotype, indicating that a different number of genes conditioning these phenotypes were segregating in each population.

### QTLs detected for each traits

We tested all 2,552 possible marker-trait associations, based on the eight phenotypes listed above, and a total of 127 loci segregating in the 24 families. The chromosomal location and subgenomic distributions of the significant marker-trait associations that were found for the target traits are shown for LGA03 and chromosome 12 in Fig. 2 (as examples), with plots for the remaining chromosomes available on-line as supplementary material (ESM Figs. S1–4). Based on this information, we estimated the number of underlying QTLs conservatively by assuming that blocks of linked markers all associated with the same trait represented a single QTL. We showed biometrical parameters (Table 2) only for the most statistically significant marker. A summary of the QTLs detected for each of the three length-related fiber properties follows.

#### *Fiber length*

A total of 20, 35, and 47 significant marker-trait associations were detected, covering 12, 22, and 29 linked groups for Lw, HVuhm, and HVsl2.5, respectively (referred to as *FL01.1-pAR792.1*; Table 2). These associations appeared to represent only 28 non-overlapping QTLs (ESM Figs. S1 and S4). Thirteen of the non-overlapping QTLs fell in the A-subgenome and 14 fell in the D-subgenome, with one unknown (locus pAR792n, a new polymorphism that did not link to known loci). For the most complex of the three phenotypes, HVsl2.5,

**Table 1** Correlation among fiber quality traits[a] measured by the AFIS and HVI instruments. Traits correlated at absolute values of 0.93 or greater are shown in bold

| | Lw | HVuhm | Ln | HVstapl | UQLw | HVsl2.5 | HVsl50 | SFCw | HVsfc | SFCn | LnCV | LwCV | HVui | HVur | HVlnCV |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Lw | 1 | | | | | | | | | | | | | | |
| Hvuhm | 0.74 | 1 | | | | | | | | | | | | | |
| Ln | **0.94** | 0.59 | 1 | | | | | | | | | | | | |
| Hvstapl | 0.77 | **0.94** | 0.61 | 1 | | | | | | | | | | | |
| UQLw | 0.92 | 0.80 | 0.75 | 0.84 | 1 | | | | | | | | | | |
| HVsl2.5 | 0.87 | 0.78 | 0.68 | 0.82 | **0.97** | 1 | | | | | | | | | |
| HVsl50 | 0.90 | 0.79 | 0.71 | 0.83 | **0.99** | **0.99** | 1 | | | | | | | | |
| SFCw | −0.80 | −0.41 | −**0.93** | −0.43 | −0.55 | −0.46 | −0.50 | 1 | | | | | | | |
| HVsfc | −0.56 | −0.45 | −0.61 | −0.48 | −0.42 | −0.36 | −0.38 | 0.60 | 1 | | | | | | |
| SFCn | −0.75 | −0.35 | −0.92 | −0.36 | −0.46 | −0.38 | −0.42 | **0.98** | 0.58 | 1 | | | | | |
| LnCV | −0.60 | −0.20 | −0.82 | −0.20 | −0.28 | −0.16 | −0.22 | **0.93** | 0.54 | **0.95** | 1 | | | | |
| LwCV | −0.54 | −0.14 | −0.79 | −0.15 | −0.21 | −0.13 | −0.17 | 0.89 | 0.51 | **0.95** | **0.97** | 1 | | | |
| Hvui | 0.59 | 0.56 | 0.60 | 0.59 | 0.48 | 0.43 | 0.45 | −0.54 | −0.77 | −0.52 | −0.48 | −0.43 | 1 | | |
| Hvur | 0.37 | 0.25 | 0.50 | 0.22 | 0.15 | 0.09 | 0.11 | −0.55 | −0.80 | −0.56 | −0.59 | −0.57 | 0.80 | 1 | |
| HVlnCV | −0.36 | −0.16 | −0.49 | −0.17 | −0.15 | −0.09 | −0.11 | 0.53 | 0.86 | 0.55 | 0.58 | 0.56 | −0.69 | −0.87 | 1 |

[a] Lw, length by weight; HVuhm, Upper half mean length; Ln, length by number; HVstapl, staple length; UQLw, upper quartile length by weight; HVsl2.5, 2.5% span length; HVsl50, 50% span length; SFCw, short fiber content by weight; HVsfc, percentage short fiber (>0.5 inch) by weight; SFCn, short fiber content by number; LnCV, coefficient of variance (CV) for length by number; LwCV, CV for length by weight; HVui, uniformity index; HVur, uniformity ratio; HVlnCV, CV% of length by number.

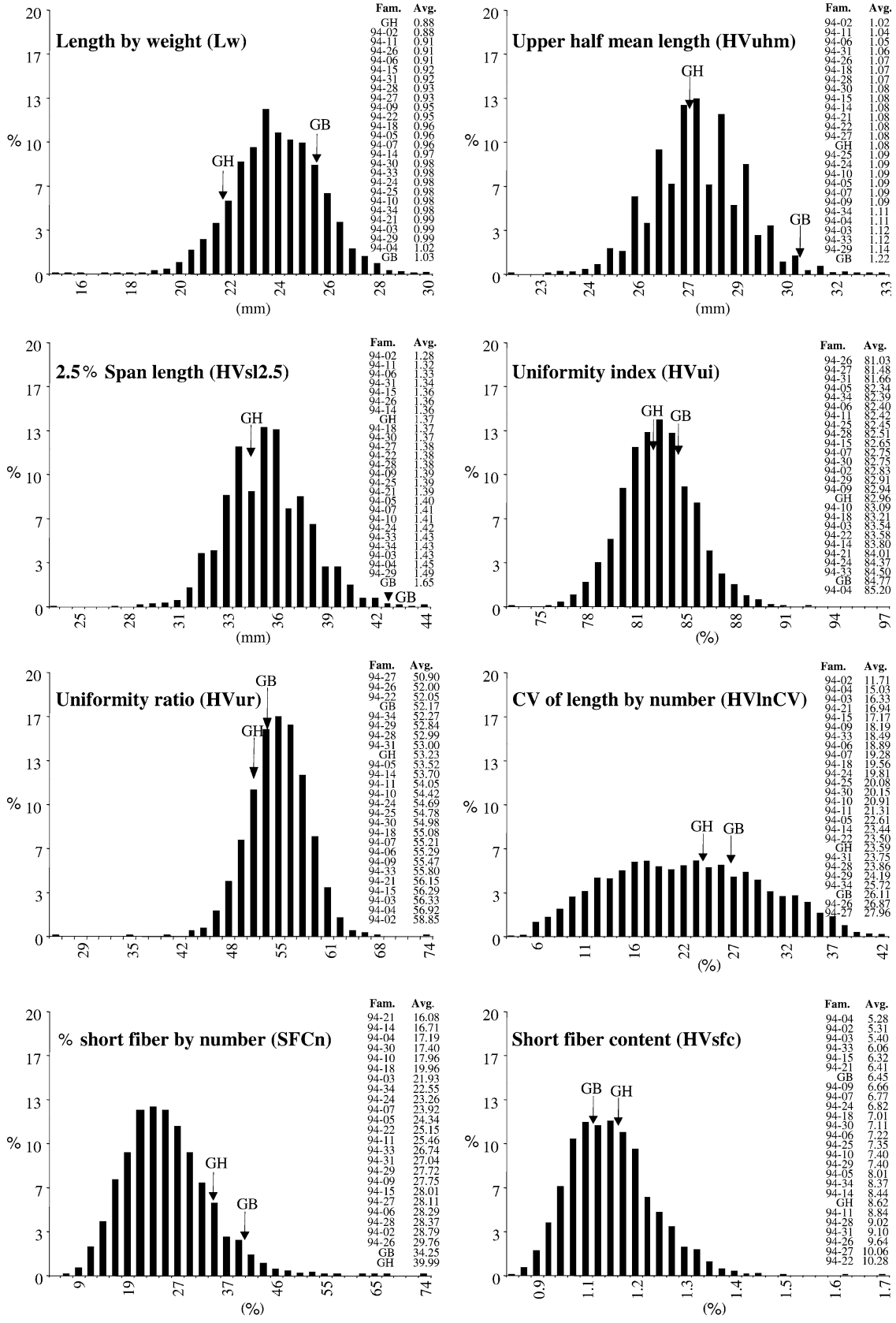



**Fig. 1** Frequency distribution of the individual $BC_3F_2$ plants and the tabulated family means of the $BC_3$ families for the eight fiber length traits

the 47 associations appeared to represent only 24 different QTLs, with four (*FL01.1*, *A03.1*, *D03.1*, and *D08.2*) reiterated in two or three $BC_3$ families. Among

the 24 HVsl2.5 QTLs, eight were also associated with both Lw and HVuhm, two with Lw only, and eight with HVuhm only. Eleven QTLs were detected for only one phenotype, including seven for HVs12.5 and four for HVUhm. The percentage of variance explained by individual associations ranged from 6% (HVuhm, *FLpAR792.1*) to 30% (Hvs12.5, *FLD08.1*). For 11 (39%) of the non-overlapping QTLs, the *G. barbadense* allele contributed increased length, which was consistent with the parental phenotypes.

### *Fiber length uniformity*

A total of five, four, and three significant marker-trait associations were detected, covering four, three, and three linked groups for HVui, HVur, and HVlnCV, respectively (*FU01.1-D08.2*; Table 2). These associations appeared to represent only nine non-overlapping QTLs (ESM Figs. S2 and S4). Six QTLs fell in the A-subgenome and three fell in the D-subgenome. For the most complex of the three phenotypes, HVui, the five associations appeared to represent four different QTLs, with none reiterated in other $BC_3$ families or by HVur or HVlnCV. Of the three QTLs each that were detected for HVur and HVlnCV, only one fell on the same chromosomal region (*FU12.2*). Variance explained by individual associations ranged from 9% (HVui, *FUD08.1*) to 24% (HVlnCV, *FU12.2*). For 6 (67%) of the non-overlapping QTLs, the allele from the favorable parent (*G. hirsutum* for HVlnCV and HVur and from *G. barbadense* for HVui) contributed increased uniformity, which was consistent with the parental phenotypes.

### *Short fiber content*

A total of five and four significant marker-trait associations were detected, covering four linkage groups each for HVsfc and SFCn. These associations appeared to represent eight non-overlapping QTLs (*SF03.1– D08.2*; Table 2). Three QTLs fell in the A-subgenome and four fell in the D-subgenome. For the more complex of the two phenotypes, HVsfc, all five associations appeared to represent different non-overlapping QTLs (*SF12.1*, *SF18.1*, *SF26.1*, *SFD08.1*, *SFD08.2*), with none reiterated in multiple $BC_3$ families. Among the five HVsfc QTLs, one (*SFD08.1*) was also associated with SFCn. Variance explained by individual associations ranged from 7% (HVsfc, *SF26.1*) to 13% (SFCn, *SF06.1*; HVsfc, *SFD08.2*). In three (38%) of the non-overlapping QTLs (*SF26.1*, *SFD08.1*, *SFD08.2*), the *G. barbadense* allele contributed reduced short fibers, which was consistent with the parental phenotypes.

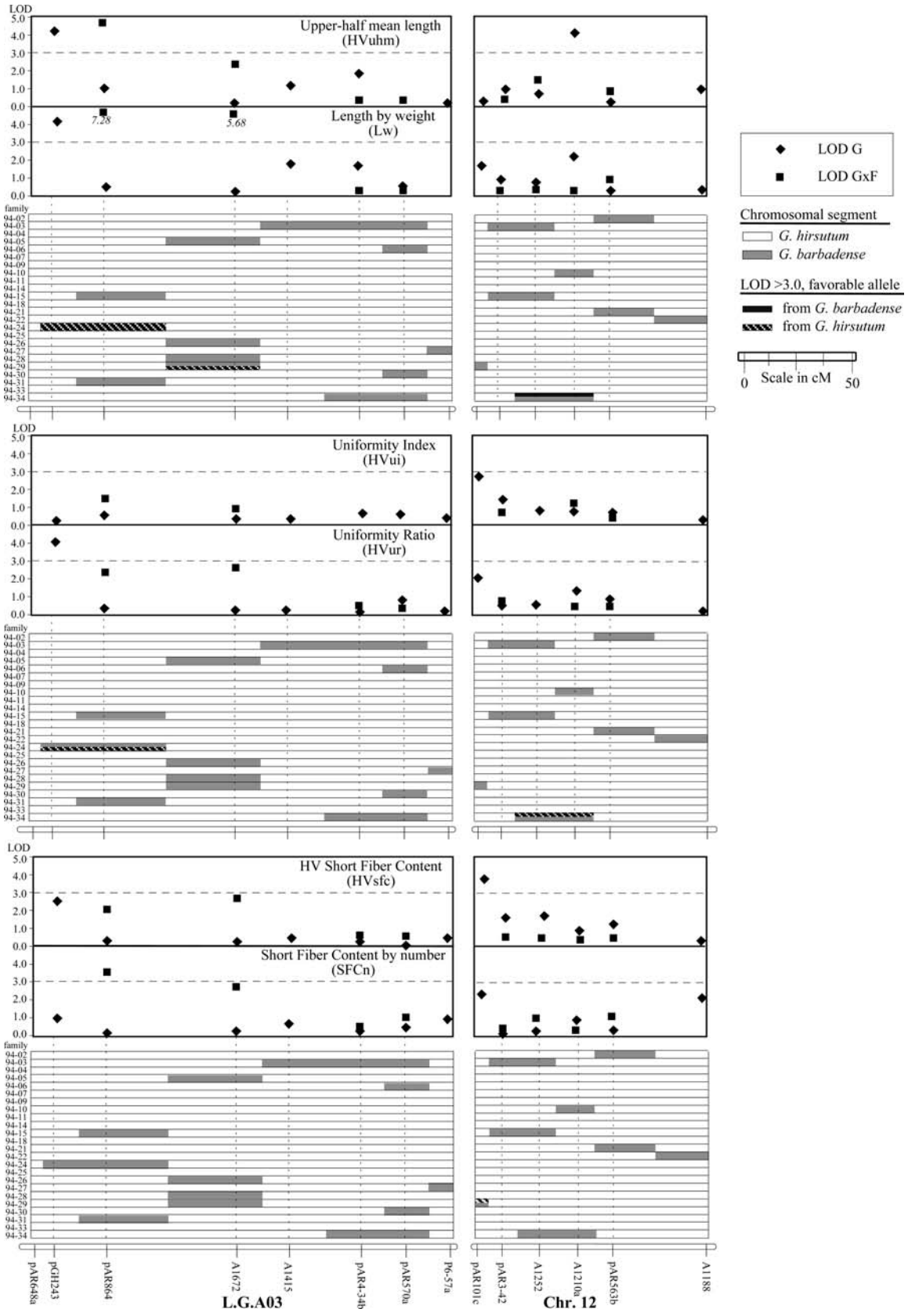
## Consistency of QTLs across families and homoeologous regions

Of the 127 loci that detected introgression, 81 linked and two unlinked loci (total of 65%) were segregating in two or more families; therefore, we performed a two-way ANOVA to test for marker-trait associations among families. Among 83 loci studied, we would expect to find less than one association that is significant at the 1% level by chance alone. A significant ($P < 0.001$) among-family G effect was detected in eight, six, and two loci for fiber length, fiber uniformity, and short fiber content, respectively. All loci showing a significant among-family G effect for fiber length (*FL01.1*, *FL03.1*, *FL04.1*, *FL07.1*, *FL12.1*, *FL14.1*, *FL14.2*, *FL23.1*) and short fiber content (*SF03.1*, *SF06.1*) also detected a QTL in the same or adjacent locus for at least one phenotype in the within-family analysis. However, for fiber uniformity, only three of the six loci significant for among-family G effects showed a QTL in the within-family analysis. Although the within-family analysis failed to detect a significant effect in any of the families segregating for these loci, the $P <$ *F* value for some families were narrowly below the $P < 0.001$ threshold for declaring a QTL (ESM Table S1).

A total of 13, two, and four loci showed G×F interactions ($P < 0.001$) for fiber length [*FL01.1* (2), *FL01.2*, *FL05.1*,*FL05.2*, *FL26.1* (2), *FLA03.1*, *FLD02.1*, *FLD03.1*, *FLD08.1*,*FLD08.2*, *pAR792n*], fiber uniformity (*FLD08.1*, *FLD08.2*), and short fiber content [*A1672* and *pAR864* (LGA03), *SFD08.1*, *SFD08.2*], respectively. All loci that showed G×F interactions for fiber length and length uniformity also detected a QTL for at least one phenotype in one family. For short fiber content, three of the four loci significant for G×F interactions showed QTLs in at least one family.

A total of 22 genomic regions were identified in which both homoeologous loci were segregating and, thus, could be tested for phenotypic associations. Fifteen regions that harbored QTLs for fiber length could be tested, and six (40%) showed QTLs for both homeologous loci (*FL01.2* and *FL15.1*; *FL05.2* and *FLD08.1*; *FL12.1* and *FL26.1*). For both length uniformity and short fiber content, five homoeologous regions could be tested for QTLs, and two regions affecting only short fiber content showed homoeologous QTLs (*SF12.1* and *SF26.1*).

**Fig. 2** Graphical genotype (*bottom section*) and significance plot of marker-trait associations (*top section*) for LGA03 and chromosome 12. For the graphical genotype, each *horizontal bar* represents the indicated chromosome for each of the 24 $BC_3F_2$ families, with *shaded areas* depicting regions containing *Gossypium barbadense* introgression. For each introgressed region, significant ($P < 0.001$) marker-trait associations detected within each family are indicated by *solid black* or *hatched bars* that represent *G. barbadense* or *G. hirsutum* favorable alleles, respectively. The *top* and *bottom half* of the *horizontal bars* correspond to significant marker-trait association for the traits indicated in the *top* and *bottom* section of the significance plot, respectively. For the significance plot, *box symbols* represent the G×F interaction and *diamond symbols* represent the G effects obtained from either among-family (two or more families) or single-family associations. *LOD*: Negative log 10 of the *P* value. By convention, cotton *chromosomes 1–13* and linkage groups with the *prefix A* are part of the A-subgenome; similarly, *chromosomes 14–26* and linkage groups with the *prefix D* are from the D-subgenome


LOD
Upper-half mean length (HVuhm)
Length by weight (Lw)
7.28
5.68
family
LOD G
LOD GxF
Chromosomal segment
G. hirsutum
G. barbadense
LOD >3.0, favorable allele
from G. barbadense
from G. hirsutum
0
Scale in cM
50
Uniformity Index (HVui)
Uniformity Ratio (HVur)
HV Short Fiber Content (HVsfc)
Short Fiber Content by number (SFCn)
pAR648a
pGH243
pAR864
A1672
A1415
pAR4-34b
pAR570b
P6-57a
pAR101c
pAR3-42
A1252
A1210a
pAR563b
A1188
L.G.A03
Chr. 12

## Discussion

The advanced backcross design permitted us to detect much larger numbers of QTLs than had previously been found to influence fiber length parameters. Within-family analysis detected 28, nine, and eight non-overlapping QTLs for fiber length, length uniformity, and short fiber content, respectively, while among-family analysis detected three additional QTLs for length uniformity. In contrast, using an interspecific $F_2$ population derived from a different *G. barbadense* by *G. hirsutum* cross, Jiang et al. (1998) detected only one QTL for fiber length and two QTLs for length uniformity. Using the $F_2$ of still another *G. barbadense* by *G. hirsutum* cross, Saranga et al. (2001) reported a similar number of QTLs (7) for length uniformity detected in the within-family analysis but only six for fiber length. The generally larger number of QTLs detected in this study as compared to $F_2$ populations support the notion that as QTLs with large effects become fixed in a population, it becomes possible to detect factors with progressively smaller effects (Paterson et al. 1990). Since many of the QTL regions detected in this study are also significant for among-family genotype effects, the validity of the QTLs is supported not only by their stringent threshold, but also by their detection in multiple genetic backgrounds.

QTL analysis of the three traits leads to several observations and conclusions with regard to the genetic control of the length of cotton fibers. First, our findings indicate that fiber length traits are governed by many minor genes with predominantly additive effects. These results are consistent with the consensus drawn from quantitative genetic variance analyses in that additive variance is the most important genetic determinant of variation in fiber length (Wilson and Wilson 1975; Tang et al. 1993; May and Green 1994). Additive QTLs are thought to be most likely to continue to function as predicted when placed in other genetic backgrounds (Tanksley and Nelson 1996; Bernacchi et al. 1998), although this remains to be tested in cotton. The large number of QTLs with small effects identified for the three length-related fiber properties (i.e., 23 QTLs detected for HVsl2.5 with the largest QTL explained less than 30% of the total phenotypic variance) reflects the general complexity of manipulating these quantitatively inherited traits. Most studies have indicated that the magnitude of genetic variation for these traits is greater than that of nongenetic factors (reviewed by May 2000), suggesting that selection should be effective in improving these traits. However, the relatively fewer QTLs for length uniformity and short fiber content suggest that the genotypic component of these traits may be less than that of fiber length. Behery (1993) showed that in addition to genetic factors that influence it directly, short fibers could also be an indirect consequence of poor fiber strength (due to fiber breakage). We will explore this possibility further in a future installment.

A second aspect of the QTL analysis that we observed is in regard to the relationships between the genes that govern the three fiber length traits. Naturally, the vast numbers of individual fibers in a sample of cotton are not all the same length. Within a single plant, fiber length variation exists among bolls from different fruiting positions (Bradow et al. 1997) and even among fibers on the same seed (Richmond and Fulton 1963). Thus, determining the genomic locations of genes conditioning fiber length, length uniformity, and short fiber content provided us an opportunity to identify regions of the genome that had effects on more than one trait. Our QTL analysis indicated that six (67%; *FU = FL for 01.1*; *03.1*; *12.1*; *A03.1*; *D08.1*; *D08.2*) and five (56%; *FU = SF for 03.1*; *06.1*; *12.1*; *D08.1*; *D08.2*) of the nine QTLs for length uniformity corresponded with QTL for fiber length and short fiber content, respectively. Similarly, four of the eight QTLs (50%; *SF = FL for 03.1*; *12.1*; *26.1*; *D08.1*; *D08.2*) for short fiber content corresponded with QTLs for fiber length. In addition, there is a near-perfect agreement between the direction of the QTL effects and the correlation among traits; an increase in fiber length would increase uniformity and decrease short fiber content. The exception is the locus on chromosome 03, where the *G. barbadense* allele is associated with increased fiber length and short fiber content but decreased fiber uniformity, contrary to expectation. Thus, co-location of many QTLs appears to account for the observed correlations among fiber length, length uniformity, and short fiber content. However, since some QTLs are unique to each trait or show opposite effects from a prediction based on the correlation among traits, this suggests that maximum genetic gain will require breeding efforts that target each trait (or an index including all three).

A third noteworthy observation from the QTL analysis concerns the contribution of the A- and D-subgenomes of allotetraploid cottons to the genetic variation of the fiber length traits. It is of much interest to determine the extent to which genetic variation in fiber length is determined by loci on the D-subgenome—i.e., from an ancestor that does not produce spinnable fiber. None of the present day D-genome diploid species, including the presumed progenitor of the AD-tetraploid cottons, produces spinnable fibers (Lee 1984; Applequist et al. 2001). Therefore, the role of the D-genome in genetic determination of fiber quality in polyploid cotton has long been a question. Over the last decade, molecular mapping studies have clearly shown that many QTLs for fiber quality map to the D-subgenome of tetraploid cotton (Jiang et al. 1998; Paterson et al. 2003; Mei et al. 2004). In fact, some studies have shown that more QTLs occur on the D-subgenome than on the A-subgenome for fiber length (Jiang et al. 1998; Paterson et al. 2003) and other traits such as resistance to bacterial blight (Wright et al. 1998) and leaf morphology (Jiang et al. 2000b). In the present study, among the 44 QTLs with known subgenome affinity that affect fiber length properties, we observed no appreciable differences in the number of QTLs between the subgenomes (22 for each

**Table 2** Biometrical parameters of QTLs affecting various fiber length phenotypes. Each row corresponds to a one-way analysis for a single locus and a single family except the three underlined loci, for which the row corresponds to a two-way analysis for multiple families

| QTL | Trait | Chromosome/ linkage group | Nearest locus | Family | $R^{2a}$ | $a^a$ | $d^a$ | $d/a$ ratio[a] | Mode of action[b] | Homeolog[c] |
|---|---|---|---|---|---|---|---|---|---|---|
| *FL01.1*[d] | Lw | Chr01 | A1686a | 94-24 | 15 | 0.03 | 0.03 | 1.00 | – | |
| *FL01.1*[d] | HVuhm | Chr01 | A1686a | 94-24 | 18 | 0.03 | 0.01 | 0.33 | A | |
| *FL01.1*[d] | HVuhm | Chr01 | A1686a | 94-29 | 17 | 0.04 | 0.00 | 0.00 | A | |
| *FL01.1*[d] | HVsl2.5 | Chr01 | A1686a | 94-24 | 24 | 0.06 | 0.02 | 0.36 | A | |
| *FL01.1*[d] | HVsl2.5 | Chr01 | A1686a | 94-26 | 12 | 0.04 | 0.00 | 0.00 | A | |
| *FL01.1*[d] | HVsl2.5 | Chr01 | A1686a | 94-29 | 16 | 0.06 | 0.02 | 0.33 | A | |
| *FL01.2* | HVsl2.5 | Chr01 | pGH468a | 94-24 | 10 | 0.031 | 0.02 | 0.60 | A | Chr15: AR077a*; pGH468b |
| *FL02.1* | Lw | Chr02 | pGH399a | 94-07 | 13 | −0.04 | 0.03 | −0.86 | A | Chr17: pGH399b |
| *FL02.1* | HVuhm | Chr02 | pGH399a | 94-07 | 12 | −0.03 | 0.02 | −0.67 | A | Chr17: pGH399b |
| *FL02.1* | HVsl2.5 | Chr02 | pGH399a | 94-07 | 18 | −0.06 | 0.03 | −0.55 | A | Chr17: pGH399b |
| *FL03.1* | HVuhm | Chr03 | pAR172b | 94-34 | 11 | 0.03 | 0.02 | 0.80 | – | Chr17: pAR172a; pAR250 |
| *FL03.1*[d] | HVsl2.5 | Chr03 | pAR172b | 94-34 | 11 | 0.05 | 0.01 | 0.20 | – | Chr17: pAR172a; pAR250 |
| *FL04.1* | HVuhm | Chr04 | A1310a | 94-34 | 12 | 0.02 | | | – | |
| *FL04.1*[d] | HVsl2.5 | Chr04 | A1310a | 94-34 | 13 | 0.03 | | | – | |
| *FL05.1* | HVsl2.5 | Chr05 | pGH530 | 94-29 | 10 | −0.05 | 0.01 | −0.22 | A | LGD08: P2-3 |
| *FL05.2* | Lw | Chr05 | G1112a | 94-29 | 8 | 0.04 | 0.02 | 0.57 | A | LGD08: G1112e; P2-9*; pAR3-41 |
| *FL05.2* | HVsl2.5 | Chr05 | G1112a | 94-03 | 9 | 0.01 | 0.05 | 10.00 | D | LGD08: G1112e; P2-9*; pAR3-41 |
| *FL06.1* | HVsl2.5 | Chr06 | A1152 | 94-26 | 0 | 0.04 | 0.01 | 0.29 | A | |
| *FL07.1*[d] | HVuhm | Chr07 | G1185a | 94-26 | 12 | −0.04 | 0.02 | −0.57 | A | Chr16: pAR564; pAR624 |
| *FL07.1* | HVsl2.5 | Chr07 | G1185a | 94-26 | 12 | −0.04 | 0.04 | −1.00 | A | Chr16: pAR564; pAR624 |
| *FL12.1*[d] | HVuhm | Chr12 | A1252 | 94-34 | 19 | 0.03 | 0.03 | 1.20 | – | Chr26: A1310a*; pGH413 |
| *FL12.1*[d] | HVsl2.5 | Chr12 | A1252 | 94-34 | 24 | 0.04 | 0.07 | 2.00 | – | Chr26: A1310a*; pGH413 |
| *FL14.1*[d] | HVuhm | Chr14 | A1727 | 94-09 | 11 | 0.03 | −0.01 | −0.40 | A | |
| *FL14.2*[d] | HVuhm | Chr14 | A1580 | 94-07 | 11 | 0.02 | 0.01 | 0.50 | – | Chr02: pAR451a |
| *FL15.1* | HVsl2.5 | Chr15 | pAR077a | 94-03 | 9 | 0.04 | −0.01 | −0.29 | A | Chr01: G1097; pGH468a* |
| *FL17.1* | HVuhm | Chr17 | pAR1-56a | 94-34 | 16 | 0.02 | 0.02 | 1.00 | A | Chr03: pAR1-56b |
| *FL17.1* | HVsl2.5 | Chr17 | pAR1-56a | 94-34 | 17 | 0.04 | 0.01 | 0.25 | A | Chr03: pAR1-56b |
| *FL20.1* | Lw | Chr20 | A1158b | 94-29 | 16 | −0.05 | 0.04 | −0.89 | A | Chr10: A1158a |
| *FL20.1* | HVsl2.5 | Chr20 | A1158b | 94-29 | 17 | −0.06 | 0.06 | −1.09 | D, A | Chr10: A1158a |
| *FL20.2* | HVsl2.5 | Chr20 | G1104 | 94-15 | 8 | 0.02 | | 0.00 | – | |
| *FL22.1* | HVuhm | Chr22 | pAR206a | 94-04 | 27 | 0.03 | 0.03 | 1.00 | – | Chr05: pAR206b |
| *FL23.1*[d] | Lw | Chr23 | pAR547 | 94-24 | 13 | 0.03 | 0.03 | 1.20 | A | |
| *FL23.1*[d] | HVuhm | Chr23 | pAR547 | 94-24 | 14 | 0.03 | 0.02 | 0.80 | A | |
| *FL23.1*[d] | HVsl2.5 | Chr23 | pAR547 | 94-24 | 16 | 0.04 | 0.04 | 1.14 | A | |
| *FL26.1* | HVuhm | Chr26 | pAR101b | 94-24 | 15 | 0.03 | −0.04 | −1.60 | – | Chr12: pAR101c; pAR3-42 |
| *FL26.1* | HVsl2.5 | Chr26 | pAR101b | 94-24 | 18 | 0.04 | −0.06 | −1.50 | A | Chr12: pAR101c; pAR3-42 |
| *FL26.1* | Lw | Chr26 | A1310b | 94-24 | 15 | 0.02 | 0.04 | 2.00 | D | Chr12: A1252* |
| *FLA01.1* | HVsl2.5 | LGA01 | G1125b | 94-28 | 10 | 0.03 | 0.02 | 0.67 | – | |
| *FLA02.1* | HVuhm | LGA02 | pAR792 | 94-15 | 9 | −0.03 | 0.01 | −0.40 | A | |
| *FLA02.1* | HVsl2.5 | LGA02 | pAR792 | 94-15 | 11 | −0.04 | −0.01 | 0.25 | A | |
| *FLA03.1* | Lw | LGA03 | A1672 | 94-29 | 9 | −0.04 | | | – | LGD02: A1174; pGH505 |
| *FLA03.1* | HVsl2.5 | LGA03 | A1672 | 94-29 | 8 | −0.04 | | | – | LGD02: A1174; pGH505 |
| *FLA03.1* | Lw | LGA03 | pAR864 | 94-24 | 26 | −0.06 | −0.03 | 0.55 | A | LGD02: pAR566a |
| *FLA03.1* | HVsl2.5 | LGA03 | pGH243 | 94-24 | 26 | −0.07 | −0.03 | 0.43 | A | LGD02: pAR566a |
| *FLA03.1* | HVuhm | LGA03 | pGH243 | 94-24 | 29 | −0.04 | −0.05 | 1.25 | D, A | LGD02: pAR566a |
| *FLD02.1* | HVuhm | LGD02 | A1413 | 94-06 | 15 | 0.07 | | | | |
| *FLD02.1* | HVuhm | LGD02 | pAR4-34a | 94-06 | 24 | −0.04 | 0.00 | 0.00 | A | LGA03: A1415; pAR4-34b |
| *FLD02.2* | HVsl2.5 | LGD02 | A1296 | 94-21 | 24 | 0.05 | −0.05 | −1.11 | A | |
| *FLD03.1* | Lw | LGD03 | pAR571b | 94-24 | 14 | 0.06 | | | – | |
| *FLD03.1* | HVuhm | LGD03 | pAR571b | 94-24 | 20 | 0.06 | | | – | |
| *FLD03.1* | HVuhm | LGD03 | pAR571b | 94-29 | 13 | 0.05 | | | – | |
| *FLD03.1* | HVsl2.5 | LGD03 | pAR571b | 94-24 | 23 | 0.09 | | | – | |
| *FLD03.1* | HVsl2.5 | LGD03 | pAR571b | 94-29 | 14 | 0.09 | | | – | |
| *FLD08.1* | Lw | LGD08 | P2-9 | 94-24 | 25 | 0.06 | −0.02 | −0.36 | A | Chr05: A1318b; G1112a* |
| *FLD08.1* | HVuhm | LGD08 | P2-9 | 94-24 | 26 | 0.05 | −0.02 | −0.40 | A | Chr05: A1318b; G1112a |
| *FLD08.1* | HVsl2.5 | LGD08 | P2-9 | 94-24 | 30 | 0.07 | −0.05 | −0.77 | A | Chr05: A1318b; G1112a* |
| *FLD08.2* | Lw | LGD08 | pAR137a | 94-28 | 11 | 0.03 | 0.00 | 0.00 | A | |
| *FLD08.2* | Lw | LGD08 | pAR137a | 94-29 | 9 | −0.03 | 0.05 | −2.00 | A | |
| *FLD08.2* | HVuhm | LGD08 | pAR137a | 94-28 | 14 | 0.02 | −0.01 | −0.50 | A | |
| *FLD08.2* | HVuhm | LGD08 | pAR137a | 94-29 | 9 | −0.01 | 0.05 | −5.00 | – | |
| *FLD08.2* | HVsl2.5 | LGD08 | pAR137a | 94-28 | 9 | 0.03 | 0.00 | 0.00 | A | |
| *FLD08.2* | HVsl2.5 | LGD08 | pAR137a | 94-29 | 9 | −0.02 | 0.10 | −5.00 | D | |
| *FlpAR792.1* | HVuhm | U | pAR792n | 94-29 | 6 | 0.02 | | | – | |
| *FlpAR792.1* | HVsl2.5 | U | pAR792n | 94-29 | 7 | 0.03 | | | – | |
| *FU01.1*[d] | HVlnCV | Chr01 | A1204 | 94-21 | 24 | 2.91 | 5.87 | 2.02 | D | |

**Table 2** (Contd.)

| QTL | Trait | Chromosome/ linkage group | Nearest locus | Family | $R^{2a}$ | $a^a$ | $d^a$ | $d/a$ ratio[a] | Mode of action[b] | Homeolog[c] |
|---|---|---|---|---|---|---|---|---|---|---|
| *FU03.1*[d] | HVur | Chr03 | pAR172b | 94-27 | 11 | −1.59 | −0.02 | 0.01 | A | Chr17: pAR172a; pAR250 |
| *FU06.1* | HVur | Chr06 | pAR988 | 94-07 | 12 | −2.08 | 0.54 | −0.26 | A | |
| *FU12.1*[d] | HVlnCV | Chr12 | pAR101c | 94-29 | 9 | 2.48 | | | – | Chr26: pAR101b |
| *FU12.2*[d] | HVlnCV | Chr12 | A1210a | 94-34 | 11 | 3.43 | −1.27 | −0.37 | A | |
| *FU12.2*[d] | HVur | Chr12 | A1252 | 94-34 | 20 | −1.74 | −0.95 | 0.55 | A | Chr26: A1310a; pGH413 |
| FU14.1[d] | HVur | Chr14 | A1580 | – | – | −0.82 | −0.77 | 0.94 | – | |
| FU17.1[d] | HVur | Chr17 | pAR1-56a | – | – | −0.13 | −0.40 | 0.38 | A | |
| *FU23.1* | HVui | Chr23 | pAR209 | 94-11 | 9 | 0.82 | −0.27 | −0.33 | A | |
| *FUA03.1* | HVui | LGA03 | pGH243 | 94-24 | 9 | −0.78 | −0.96 | 1.23 | A | LGD02: pAR566A |
| FUD03.1[d] | HVui | LGD03 | pAR571b | – | – | −0.33 | 0.98 | −2.99 | – | |
| *FUD08.1* | HVui | LGD08 | P2-9 | 94-24 | 8 | 1.10 | 0.08 | 0.07 | A | Chr05: A1318b; G1112a |
| *FUD08.2* | HVui | LGD08 | pAR137a | 94-28 | 12 | 0.92 | 0.25 | 0.27 | A | |
| *SF03.1*[d] | SFCn | Chr03 | pAR172b | 94-27 | 11 | 3.92 | −0.68 | −0.17 | A | Chr17: pAR172a; pAR250 |
| *SF06.1*[d] | SFCn | Chr06 | pAR988 | 94-07 | 13 | 3.24 | −0.61 | −0.19 | A | |
| *SF12.1* | HVsfc | Chr12 | pAR101c | 94-29 | 7 | 0.88 | | | – | Chr26: pAR101b* |
| *SF18.1* | HVsfc | Chr18 | P9-53b | 94-29 | 8 | 0.81 | | | – | |
| *SF20.1* | SFCn | Chr20 | A1158b | 94-29 | 9 | 2.88 | −2.43 | −0.84 | – | Chr10: A1158a |
| *SF26.1* | HVsfc | Chr26 | pAR101b | 94-29 | 8 | −0.95 | −0.68 | 0.72 | – | Chr12: pAR101c*; pAR3-42 |
| *SFD08.1* | HVsfc | LGD08 | P2-9 | 94-24 | 9 | −1.54 | −0.37 | 0.24 | A | Chr05: A1318b; G1112a |
| *SFD08.1* | SFCn | LGD08 | P2-9 | 94-24 | 9 | −1.06 | −2.41 | 2.27 | A | Chr05: A1318b; G1112a |
| *SFD08.2* | HVsfc | LGD08 | pAR137a | 94-28 | 13 | −1.27 | −0.52 | 0.41 | A | |

[a] Quantitative parameters: $R^2$, percentage of phenotypic variation explained by the marker genotype at the corresponding marker and family (missing where a significant association was not detected); *a*, additive; *d*, dominance; *d*/*a* ratio, overdominance effect

[b] Modes of gene action are indicated by: A, additivity; D, dominance; H, overdominance, with missing values corresponding to dominant or severely distorted DNA marker loci

[c] Homeologous loci, if present, are always shown, with * indicating that at least one family segregating at the homeologous loci gave a significant genotype effect

[d] QTLs in which the RFLP locus or a nearby locus showed a significant among-family effect

of the A- and D-subgenomes). However, our data add further evidence to increasingly common reports of the D-subgenome's contribution to the genetic control of fiber traits. Collectively, these results demonstrate that polyploidization in *Gossypium* appears to have created novel variation in a host of phenotypes (Wendel 2000), including fiber quality. Recruitment of D-genome genes by allopolyploidization may have permitted additional flexibility for artificial selection through breeding to develop cultivated tetraploids with a fiber length beyond that attained by the highest quality cultivated diploid cottons (Applequist et al. 2001).

While some degree of transgressive segregation in fiber length continues to be discovered from crossing divergent medium-staple parents, the use of interspecific gene combinations appears to offer an important means by which to improve this valuable trait. Although barriers to gene introgression from *G. barbadense* to *G. hirsutum* may exist (Stephens 1949; Jiang et al. 2000a), the availability of DNA markers linked to *G. barbadense* QTLs identified in this and other studies (Jiang et al. 1998; Paterson et al. 2003) could assist breeders in transferring and maintaining these traits during cultivar development. Further, because many of the QTLs from *G. barbadense* are now in a near-isogenic state in the AB-QTL population, the phenotypic effect measured for each QTL is likely to be a better predictor of its ultimate effect when transferred to other cultivated backgrounds. Finally, while QTL mapping involving interspecific crosses reported in the literature have focused entirely on *G. barbadense*, we note that three additional tetraploid *Gossypium* species are cross compatible with Upland cotton and could offer yet additional sources of allelic variation to the upland cotton germplasm.

**Acknowledgements** We acknowledge the financial support from the Texas and Georgia Agricultural Experiment Stations, Texas Higher Education Coordinating Board, Cotton Incorporated, and USDA-IFAFS. XD was a postdoctoral research associate of the Fonds National Belge de la Recherche Scientifique.

*Conflict of interest:* No information supplied

## References

Applequist WL, Cronn, Wendel JF (2001) Comparative development of fiber in wild and cultivated cotton. Evol Dev 33:3–17

Behery HM (1993) Short-fiber content and uniformity index in cotton. International Cotton Advisory Committee and Center for Agriculture and Biosciences Review Article 4. Washington, D.C., pp 1–40

Bernacchi D, Beck-Bunn T, Emmatty D, Eshed Y, Inai S, Lopez J, Petiard V, Sayama H, Uhlig J, Zamir D, Tanksley S (1998) Advanced backcross QTL analysis of tomato. II. Evaluation of near-isogenic lines carrying single-donor introgressions for desirable wild QTL-alleles derived from *Lycopersicon hirsutum* and *L. pimpinellifolium*. Theor Appl Genet 97:170–180

Bradow JM, Wartelle LY, Bauer PJ, Sassenrath-Cole GF (1997) Small-sample cotton fiber quality quantitation. J Cotton Sci 1:48–60

Chee P, Draye X, Jiang C, Decanini L, Delmonte T, Bredhauer B, Smith CW, Paterson AH (2005) Molecular dissection of interspecific variation between *Gossypium hirsutum* and *G. barbadense* (cotton) by a backcross-self approach: I. Fiber elongation (in press) doi: 10.1007/s00122-005-2063-z

Jiang C, Wright R, El-Zik K, Paterson A (1998) Polyploid formation created unique avenues for response to selection in *Gossypium* (Cotton). Proc Natl Acad Sci USA 95:4419–4424

Jiang C, Chee P, Draye X, Morrell P, Smith C, Paterson A (2000a) Multi-locus interactions restrict gene flow in advanced-generation interspecific populations of polyploid *Gossypium* (Cotton). Evolution 54:798–814

Jiang C, Wright R, Woo S, Delmonte T, Paterson A (2000b) QTL analysis of leaf morphology in tetraploid *Gossypium* (Cotton). Theor Appl Genet 100:409–418

Jost P (2002) Planning for fiber quality. In: Jose P (ed) UGA Georgia Cotton Newsletter. http://www.griffin.peachnet.edu/caes/cotton/cnl041902.htm

Lee JA (1984) Cotton as a world crop. In: Kohel RJ, Lewis CF (eds) Cotton. American Society of Agronomy, Madison, pp 6–24

May OL (2000) Genetic variation for fiber quality. In: Basra AS (ed) Cotton fibers–developmental biology, quality improvement, and textile processing. Food Products Press, New York, pp 183–229

May OL, Green CC (1994) Genetic variation for fiber properties in elite Pee Dee cotton populations. Crop Sci 34:684–690

Mei M, Syed NH, Gao W, Thaxton PM, Smith CW, Stelly DM, Chen ZJ (2004) Genetic mapping and QTL analysis of fiber-related traits in cotton (*Gossypium*). Theor Appl Genet 108:280–291

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

Paterson AH, Saranga Y, Menz M, Jiang C, Wright RJ (2003) QTL analysis of genotype×environmental interactions affecting cotton fiber quality. Theor Appl Genet 106:384–396

Perkins HH, Ethridge DE, Bragg CK (1984) Fiber. In: Kohel RJ, Lewis CF (eds) Cotton. American Society of Agronomy, Madison, pp 438–508

Reinisch AJ, Dong JM, Brubaker C, Stelly D, Wendel JF, Paterson AP (1994) A detailed RFLP map of cotton (*Gossypium hirsutum*×*Gossypium barbadense*): chromosome organization and evolution in a disomic polyploid genome. Genetics 138:829–847

Richmond TR, Fulton HJ (1963) Variability of fiber length in a relatively uniform strain of cotton. J Agric Res 35:749–763

Rong J, Abbey C, Bowers JE, Brubaker CL, Chang C, Chee PW, Delmonte TA, Ding X, Garza JJ, Park BSMC-h, Pierce GJ, Rainey KM, Rastogi VK, Schulze SR, Trolinder NL, Wendel JF, Wilkins TA, Williams-Coplin TK, Wing RA, Wright RJ, Zhao X, Zhu L, Paterson AH (2004) A 3,347-locus genetic recombination map of sequence-tagged sites reveals features of genome organization, transmission and evolution of cotton (*Gossypium*). Genetics 166:389–417

Saranga Y, Menz M, Jiang C, Wright R, Yakir D, Paterson AH (2001) Genomic dissection of genotype×environment interactions conferring adaptation of cotton to arid conditions. Genome Res 11:1988–1995

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

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

Steadman RG (1997) Cotton testing. In: Smirfitt JA (ed) Textile progress. The Textile Institute, Manchester, pp 15–23

Stephens SG (1949) The cytogenetics of speciation in *Gossypium*. I. Selective elimination of the donor parent genotype in interspecific backcrosses. Genetics 34:627–637

Tang B, Jenkins JN, McCarty JC, Watson CE (1993) $F_2$ hybrids of host plant germplasm and cotton cultivars. II. Heterosis and combining ability for fiber properties. Crop Sci 33:706–710

Tanksley SD, Nelson CJ (1996) Advanced backcross QTL analysis: a method for the simultaneously discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. Theor Appl Genet 92:191–203

Wendel JF (2000) Genome evolution in polyploids. Plant Mol Biol 2:225–249

Wilson FD, Wilson RL (1975) Breeding potentials of noncultivated cottons. I. Some agronomic and fiber properties of selected parents and their $F_1$ hybrids. Crop Sci 15:763–766

Wright R, Thaxton P, El-Zik K, Paterson A (1998) D-subgenome bias of Xcm resistance genes in tetraploid *Gossypium* (Cotton) suggests that polyploid formation has created novel avenues for evolution. Genetics 149:1987–1996

Young ND, Tanksley SD (1989) RFLP analysis of the size of chromosomal segments retained around the *Tm-2* locus of tomato during backcross breeding. Theor Appl Genet 77:353–359