Theor Appl Genet (2005) 111: 1137–1146
DOI 10.1007/s00122-005-0041-0

ORIGINAL PAPER

**Junkang Rong · Gary J. Pierce · Vijay N. Waghmare
Carl J. Rogers · Aparna Desai · Peng W. Chee
O. Lloyd May · John R. Gannaway
Jonathan F. Wendel · Thea A. Wilkins
Andrew H. Paterson**

Genetic mapping and comparative analysis of seven mutants related to seed fiber development in cotton

Received: 18 April 2005 / Accepted: 2 July 2005 / Published online: 2 August 2005

Abstract Mapping of genes that play major roles in cotton fiber development is an important step toward their cloning and manipulation, and provides a test of their relationships (if any) to agriculturally-important QTLs. Seven previously identified fiber mutants, four dominant (Li_1, Li_2, N_1 and *Fbl*) and three recessive (n_2, *sma-4*(h_a), and *sma-4(fz)*), were genetically mapped in six F_2 populations comprising 124 or more plants each. For those mutants previously assigned to chromosomes by using aneuploids or by linkage to other morphological markers, all map locations were concordant except n_2, which mapped to the homoeolog of the chromosome previously reported. Three mutations with primary effects on fuzz fibers (N_1, *Fbl*, n_2) mapped near the likelihood peaks for QTLs that affected lint fiber productivity in the same populations, perhaps suggesting pleiotropic effects on both fiber types. However, only Li_1 mapped within the likelihood interval for 191 previously detected lint fiber QTLs discovered in non-mutant crosses, suggesting that these mutations may occur in genes that played early roles in cotton fiber evolution, and for which new allelic variants are quickly eliminated from improved germplasm. A close positional association between *sma-4*(h_a), two leaf and stem-borne trichome mutants (t_1, t_2), and a gene previously implicated in fiber development, sucrose synthase, raises questions about the possibility that these genes may be functionally related. Increasing knowledge of the correspondence of the cotton and *Arabidopsis* genomes provides several avenues by which genetic dissection of cotton fiber development may be accelerated.

Communicated by F. Salamini

J. Rong · G. J. Pierce · V. N. Waghmare · C. J. Rogers
A. Desai · P. W. Chee · A. H. Paterson (✉)
Plant Genome Mapping Laboratory,
University of Georgia, Athens, GA 30602, USA

A. Desai · A. H. Paterson
Department of Crop and Soil Sciences,
University of Georgia,
3111 Miller Plant Sciences Building,
Athens, GA 30602, USA

O. L. May · P. W. Chee
Coastal Plains Experiment Station,
University of Georgia, Tifton, GA 31793, USA

J. R. Gannaway
Texas A&M Research and Extension Center,
Lubbock, TX 79401, USA

J. F. Wendel
Department of Ecology, Evolution,
and Organismal Biology, Iowa State University,
Ames, IA 50011, USA

T. A. Wilkins
Department of Plant Sciences, University of California,
Davis, CA 95616, USA

Introduction

A few members of the *Gossypium* (cotton) genus are cultivated for the production of fibers that make possible world cotton commerce of about $20 billion annually (considering only the value of raw material). Anatomically, cotton seed fibers are single-celled trichomes that differentiate from the outermost cell layer (protoderm) of the ovule. Fiber development occurs in four discrete, yet overlapping stages - differentiation, expansion, secondary cell wall synthesis, and maturation. The cellular activities that characterize each stage of fiber morphogenesis are governed by master switches in the developmental program that result in dynamic changes in the cotton fiber transcriptome (Wilkins and Jernstedt 1999; Arpat et al. 2004; Wilkins and Arpat 2005; Wilkins et al. 2005)

The seed of all 50 *Gossypium* species have epidermal hairs, and in a few species may contribute to saltwater dispersal (Fryxell 1979). Evolution of the economically-

important cotton 'lint' fiber is thought to have occurred in four stages:

1. After its divergence from a common ancestor shared with the F genome about 5 million years ago (MYA, (Cronn et al. 2002), natural selection in the A-genome lineage is thought to have differentiated these fibers into two distinct types, 'fuzz' and 'lint' (Fryxell 1963). Fuzz fibers are short (< 5 mm) and tightly appressed to the seed coat, similar to the seed hairs of wild cottons. Economically-important lint fibers fall into three classes based on average length ("staple") at maturity: short staple (< 21 mm), *G. arboreum* and *G. herbaceum*; long staple (28–32 mm), *G. hirsutum* ('upland'); and extra-long staple (> 35 mm), *G. barbadense* ('Pima', 'Egyptian', or 'Sea Island').
2. Polyploidization about 1 MYA (Wendel 1989), joined in a common nucleus the divergent A- and D-genomes. Wild tetraploid cottons and their A genome diploid ancestors have lint fibers while D genome diploids do not.
3. Domestication, independently in the A-genome diploids *G. herbaceum* and *G. arboreum*, and the tetraploids *G. hirsutum* and *G. barbadense*, substantially improved cotton fiber yield and quality. Domesticated A-genome diploids may have existed in the Old World as early as 8000 years ago (Moulherat et al. 2002), and domesticated tetraploid cottons existed in the New World by 5500–4300 years ago (Stephens and Moseley 1974).
4. Scientific improvement of *G. hirsutum* (AD1) has emphasized maximum yield, while *G. barbadense* (AD2) is bred for fibers of superior length, strength, and fineness. *G. herbaceum* and *G. arboreum* are bred and cultivated in Asia for high stress tolerance that partly compensates for generally lower quality and yield than the tetraploids.

Little is known about the specific genes responsible for variation in cotton fiber quality, although a growing set of clues exist from which to identify such genes. At least 191 quantitative trait loci (QTLs) have been genetically mapped that contribute to attributes which differentiate between *G. hirsutum* and *G. barbadense*(Jiang et al. 1998; Saranga et al. 2001; Paterson et al. 2003; Chee et al. 2005a; Chee et al. 2005b; Draye et al. 2005), and even among elite *G. hirsutum* genotypes in a few cases (Zhang et al. 2003). A long history of research provides information about the developmental timetable associated with cotton fiber initiation, elongation, and maturation. Recently the evolutionary timetable was clarified (Basra and Malik 1984; Applequist et al. 2001) and growing information exists about expression patterns of candidate genes (Turley and Ferguson 1996; Wang et al. 2001; Arpat et al. 2004). The genetic complexity of the cotton fiber transcriptome accounts for as much as 50% of the entire cotton transcriptome. Although ~40% of fiber genes encode unknown functions, about 26% of the genes are developmentally regulated and are primary targets for genetic mapping and functional analysis (Arpat et al. 2004; Rong et al. 2004; Wilkins and Arpat 2005; Wilkins et al. 2005). Recent work implicates sucrose synthase (*Sus*) as one potential determinant of fiber initiation and elongation (Ruan et al. 1998; Ruan et al. 2003) as RNAi suppression in transgenic plants produces a phenotype resembling those of non-fiber-producing cotton genotypes. An integrative approach to identify specific genes and underlying genetic mechanisms responsible for the evolution of this trait is timely.

Mutants are a powerful resource for studying gene function. In cotton, several mutants for fiber development have been discovered and some have been assigned to cytologically-identifiable chromosomes using aneuploid stocks (Endrizzi et al. 1985; Percy and Kohel 1999; Karaca et al. 2002; Kohel et al. 2002). Among them, *ligon lintless-1* (Karaca et al. 2002) and *ligon lintless-2* (Li_2) were reported to be monogenic and dominant, causing extreme reductions in lint fiber length (to less than 10 mm) on mature seeds (Griffee and Ligon 1929; Kohel et al. 1992). N_1 and n_2 result in a naked seed phenotype (Percy and Kohel 1999). *Fbl* is an incompletely dominant fiberless mutation exhibiting no lint or fuzz fibers (Kearney and Harrison 1927). A genetic stock, SMA-4, contains a recessive mutation (h_a) that confers fiberless seed, and an epistatic recessive mutation (*fz*) that produces lintless (i.e. fuzz fibers only) seed in the absence of homozygosity for h_a. These phenotypes appear to be temperature dependent when unfertilized ovules are cultured in vitro (Beasley and Egli 1977).

Genetic mapping of mutants is a valuable step toward their isolation, and may also provide other clues to their organization and function. So far, only Li_1 has been roughly mapped on Chr. 22 with 4 SSR markers (Karaca et al. 2002). We sought to locate seven fiber mutants to small (ca. 5 cM) intervals of the cotton chromosomes using a detailed map of sequence-tagged-site based loci (Rong et al 2004), and to use this information to evaluate their relationships to one another and to other traits including QTLs previously mapped in interspecific crosses between elite cottons.

Materials and methods

Mapping populations

The symbols, phenotypes and stocks of seven fiber mutants studied herein were described in Table 1. All field populations were grown at the Coastal Plain Experiment Station, Tifton, GA.

Phenotype analysis

The seed cotton on F_2 plants was harvested at three dates each about 2 weeks apart, with fiber for individual

Table 1 Fiber mutants and their characteristics

Stock[1]	Gene	Phenotype	Reference/Source
SMA-4	*sma-4(h$_a$), sma-4(fz)*	Glabrous lintless seed (h$_a$), or fuzz-fibered seed (*H$_a$-fzfz*); temperature-dependent in vitro	Beasley and Egli 1977
Fiberless	*Fbl*	Large-seeded, no lint or fuzz fibers (fiberless), incompletely dominant	J.McD. Stewart, University of Arkansas
Naked seed	N_1	No fuzz fibers	USDA Collection
Naked seed	n_2	No fuzz fibers	USDA Collection
Ligon-lintless 1	Li_1	Shortened lint fibers (≤ 5 mm), deformed leaves and stems, stunted	USDA Collection
Ligon-lintless 2	Li_2	Shortened lint fibers (≤ 5 mm)	USDA Collection; Kohel et al 1992

[1] All stocks are *G. hirsutum*, with exception of SMA-4 (*G. arboreum*), provided by J. Wendel; all tetraploid stocks were provided by T. Wilkins. The mutants indicated as being from the USDA collection were provided to T. Wilkins by courtesy of Russell Kohel.

plants pooled across dates and phenotyped in the laboratory. Figure 1 illustrates the qualitative phenotypes, as detailed in the text below.

Segregation of the qualitative phenotypes was analyzed with chi-squared tests. In several cases, the mutations had qualitative effects on fuzz fibers, however we observed quantitative variation in lint fiber characteristics as well. To explore this, for the tetraploid populations, Lint % in seed cotton was measured by subtracting seed weight from the weight of seed cotton (before ginning), then dividing by the weight of seed cotton and multiplying by 100. Lint Index was calculated by subtracting delinted seed weight (in grams) from the weight of seed cotton of 100 seeds. The correlations among these traits were analyzed using Excel.

DNA extraction and map construction

DNA extraction, restriction digestion, gel electrophoresis, Southern blotting, probe labeling, Southern hybridization, and autoradiography was previously described (Reinisch et al. 1994). Blots including Pima S-7, 9 wild-type and 9 mutant plants were first prepared for each of 4 (Li_1, Li_2, N_1 and *Fbl*) F_2 mapping populations to screen for associations of mutant genes with DNA

Fig. 1a–k Wild and mutant seed cotton. **a** Wild type, both fuzz and lint fibers present in all tetraploid populations; **b** Fuzz-fibered but lintless in Li_1; **c** Fuzz-fibered but lintless in Li_2; **d** Sparsely linted in n_2; **e** Naked with lint in n_2; **f** Naked with lint in N_1; **g** Naked without lint in N_1; **h** Naked with lint in *Fbl*; **i** Naked without lint in Fbl; **j** Wild type, both fuzz and lint fibers present in SMA-4; **k** Naked without lint or tufted in SMA-4

markers selected at ca. 20 cM intervals from a detailed map (Rong et al. 2004). Tentative associations with specific DNA markers were then further tested by mapping additional nearby markers on all F_2 plants in the appropriate cross. The name and origins of all genomic or cDNA probes were described elsewhere (Rong et al. 2004).

Linkage analysis was performed using Mapmaker 3.0 (Lander et al. 1987). Initial groups of linked markers were identified using the "group" command, then marker orders refined with commands of "compare" and "try" and finally tested using "ripple". Genetic distances in centiMorgans (cM) were calculated using the Kosambi mapping function (Kosambi 1944). QTL likelihood maps, gene action, and phenotypic variance explained by individual QTL were determined by interval mapping using Mapmaker/QTL. A stringent logarithm of odds ratio (LOD) threshold of >3.0 was used to determine the position of QTLs, appropriate for the large recombinational length and number of chromosomes in the cotton genome (Lander and Botstein 1989).

Results

Phenotypes and segregation of fiber mutation

Fiber phenotypes conferred by the seven mutants are shown in Fig. 1. For all genotypes, the wild type phenotype had both fuzz and lint fibers (Fig. 1a for tetraploid, j for diploid). For Li_1 and Li_2, the mutant had fibers shorter than 10 mm or absent (Fig. 1b, c), and lint and fuzz fibers were not clearly distinguishable. For N_1, *Fbl*, and *sma*, the mutant was naked, completely lacking fuzz fibers with (Fig. 1f and h) or without (Fig. 1g, i and k) lint fibers. For n_2, the mutant produces lint, but bears a naked seed phenotype that differs slightly from N_1, with distal capped tufts of fuzz fibers present at the micropyle and chalazal tips of the seed (Fig. 1d and e). For the cross involving SMA-4, the linted versus lintless and fuzz-fibered versus naked phenotypes were both scored.

Consistent with classical reports (Narbuth and Kohel 1990), the Li_1 mutant was not only lintless but also exhibited deformed leaves and stems and was somewhat stunted in its growth, while Li_2 was free of these traits. Li_1 also had somewhat larger seeds (Fig. 1b versus c), perhaps as a result of maturing fewer bolls. *N1* and *Fbl* had similar fiber phenotypes and could be easily classified into naked (with, Fig. 1f, h or without lint, Fig. 1g, i) versus fuzzy cotton seeds. Naked seeds showed much variation in lint production (Fig. 1g and i). The F_2 plants of n_2 and *sma-4* were less easily classified, with some seeds appearing 'tufted,' or fiberless (*sma-4*, Fig. 1k) or sparsely linted (n_2, Fig. 1d).

To map five fiber mutants found in tetraploid cotton (Table 1), *G. barbadense* L. cv. Pima S-7 was crossed as a common parent with five *G. hirsutum* stocks each carrying a specific mutation. To map two mutants in the diploid *G. arboreum* accession SMA-4 (PI529740), the accession was crossed with *G. herbaceum* accession A_1-97 (PI529670). Hybridity of the resulting F_1 seeds was confirmed using DNA markers. F_1 plants were selfed to produce F_2 seeds in the greenhouse. The following numbers of F_2 plants were developed for individual mapping populations: 136 for Li_1, 154 for Li_2, 143 for N_1, 140 for *Fbl*, 124 for n_2 and 167 for *SMA-4*.

The segregation of fiber phenotypes was investigated for all F_2 populations (Table 2). Chi-squared tests indicated that fiber segregation (fuzzy to naked) in N_1 and *Fbl* populations completely fit single dominant gene models, and alternative models could be excluded (data not shown). Segregation at the Li_2 locus favored the wild type, although the observed ratio did not significantly ($p < 0.05$) deviate from a single dominant-gene model. Segregation of Li_1 also favored the wild type, deviating significantly from a single dominant-gene model ($P = 1.1\times10^{-7}$) presumably due to the failure of many homozygous mutants (which had deformed stems and leaves with stunted plants) to produce fiber or even survive. The sparsely linted seeds of some n_2 plants complicated analysis – if these were considered linted, then segregation does not deviate significantly from a single recessive gene model (98:25, Table 2). However, if the sparsely linted seeds are considered naked, wild-type and mutant are about equally frequent, suggesting a large deviation from single-gene inheritance. The *SMA-4* cross segregated for two phenotypes, naked and fiberless, that were analyzed separately. The segregation of fiberless fit the 3:1 ratio expected for a single recessive gene (Table 2). However, the naked phenotype deviated from the one recessive gene model ($P = 7.81\times10^{-5}$, Table 2), containing far too many mutants. Classical studies (Silow 1941) have suggested that the homozygous h_a allele (sma-4 fiberless) is epistatic to the *fz* allele, resulting in a mutant phenotype regardless of genotype at the *fz* locus. Consistent with this, segregation of the naked phenotype for this stock did not deviate significantly from a 9:7 ratio.

Table 2 Chi-squared test for the segregation of cotton fiber phenotype caused by mutant genes

F_2 Population	Total	W[a]	M[b]	Pr (3:1)
Pima S-7× Li_1	151	66	85	1.1×10^{-7}
Pima S-7× Li_2	158	50	108	0.054
Pima S-7× N_1	159	39	120	0.891
Pima S-7× *Fbl*	153	31	122	0.176
Pima S-7× n_2	123	98	25	0.231
SMA4×A1-97 (naked)[c]	158	97	61	7.8×10^{-5}
SMA4×A1-97 (fiberless)	158	112	46	0.232

a: number of wild type (W) plants; Fuzzy with lint. In n_2, the plants with sparsely linted seeds were included as wild type.
b: number of mutant (M) plants; Fuzzy without lint in Li_1 and Li_2; naked with or without lint in N_1, *Fbl*,*sma-4(naked)*; naked with lint in n_2; fiberless in *sma-4*(h_a).
c: As detailed in text, this phenotype is thought to have an epistatic relationship with the fiberless phenotype. Based on this, we also evaluated a two-gene dominant model (9:7), and found non-significant deviation (0.193).

It is noteworthy that Pima S7, our 'wild type' parent, can be argued to have a form of naked seed. The lint fibers do attach to the seed coat, albeit delicately. The fuzz fibers do not attach to the seed coat, except for distal capped tufts resembling the n_2 mutant. In principle, one might expect this to segregate in crosses to *G. hirsutum* genotypes containing the various mutants. However, for the Li_1 and Li_2 populations, together with a third population that showed no phenotype and consequently was not mapped, we found only 2 (of about 400) plants for which seed were considered naked, suggesting that the Pima S7 phenotype may be complex. Consequently, while this phenotype of Pima S7 warrants further study, it does not appear to interfere with our mapping of other mutants.

Genetic mapping of qualitative fiber mutations

To plot these loci relative to a detailed reference molecular map (Rong et al. 2004), we used the two-step process described (see Methods), first screening a small subset of plants with markers at ca. 20 cM intervals, then adding more markers in regions that show association. The details of mapping each mutant are as follows.

Li_1 The probe Gate4AD10 showed no recombination with Li_1 in the survey population. Detailed mapping of the full population revealed Li_1 to be on Chr. 22, consistent with a prior report (Karaca et al. 2002). We mapped in detail a chromosome segment that corresponds to the region from 19.5–59.9 cM of the reference map of Chr. 22 (Rong et al. 2004), with virtually identical marker order except for one locus (pAR0711) that appears to have revealed polymorphism at different paralogous loci in the two populations. The new map is somewhat longer than the published one (55.8 vs 40 cM). The region containing Li_1 shows unusually high marker density in the reference map, and has tentatively been suggested to include the centromere (Rong et al. 2004). The closest markers flanking Li_1 are Gate4CA09 and Coau1J04, 2.7 and 1.3 cM away, respectively.

Li_2 The probe pAR0947 co-segregated with the Li_2 fiber phenotype. On the reference map (Rong et al. 2004), pAR0947 was mapped to the upper region of Chr. 18 (Fig. 2), consistent with assignment of Li_2 based on cytogenetic methods (Kohel et al. 2002). A map including 15 RFLP loci and Li_2 spans 82.1 cM (Fig. 2), about 22 cM longer than the reference map of this region. Li_2 is located near the terminus of the map (and presumably the chromosome), 0.5 cM from A1552 and 1.8 cM from two co-segregating markers, Gate4BF10 and Coau1O05. Each of the 5 markers nearest to Li_2 is dominant or recessive, possibly masking recombinants in some F_2 plants and slightly reducing the precision of our estimates of genetic distance in the region.

n_2 Due to uncertainty raised by the sparsely linted phenotype and concern about the possible involvement of multiple loci, the entire n_2 population was mapped with all survey markers, at about 20 cM intervals. The n_2 locus had been previously assigned to Chr. 26 using aneuploids (Endrizzi and Ramsay 1980). Our map of Chr. 26 in the n_2 cross included 18 markers and spanned 191 cM, lacking only the terminal 7.2 cM of the reference map (Rong et al. 2004). However, no markers linked to n_2 were found on Chr. 26, regardless of whether the sparsely linted seeds were treated as missing data, wild type (fibered) or mutant (fuzzless or naked seed). However, n_2 was linked to several markers from the homoeolog, Chr. 12, when the sparsely linted seeds were scored as missing data (linked to 10 loci, LOD > 3.0), wild type (10 loci) or mutant (4 loci). We further enriched the Chr. 12 map to include 33 RFLP loci and span 163.4 cM, 23.3 cM shorter than the corresponding region of the reference map (Fig. 2). When the sparsely linted seeds were scored as missing data or wild type, n_2 mapped in a region of 27.9 cM flanked by the markers pAR0155 and Gate2BB08, 13.3 cM and 14.6 cM away, respectively. When the n_2 locus was excluded from the map, the interval between pAR0155 and Gate2BB08 decreased from 27.9 cM to 17.0 cM. This may suggest that either the n_2 phenotype has incomplete expressivity and/or imperfect penetrance, and therefore some phenotypic assignments conflict with flanking DNA marker genotypes. QTL analysis (detailed below) based on Lint % and Lint Index strongly supported the location of n_2 in this general region of Chr. 12.

N_1 and Fbl The same survey probe, Unig24G11, showed only one recombinant with both N_1 and *Fbl*. On the reference map (Rong et al. 2004), Unig24G11 was mapped on Chr. 12 (Fig. 2), consistent with prior assignment of N_1 (Kohel 1979). Detailed genetic maps of Chr. 12 for the N_1 and *Fbl* populations, respectively (Fig. 2) were very similar to the reference map and to one another, differing only in that 1) the *Fbl* map was 17 cM longer than N_1, 2) pAR10F10 and pAR0155 could be mapped on N_1 only, and 3) Gate2BD04, pAR0244 and Gate1AC08 could be mapped on *Fbl* only. Most of the markers on these two maps also shared the same linear order on the n_2 map. N_1 and *Fbl* were both mapped to the same region of about 5 cM in the reference map of Chr. 12. N_1 was flanked on one side by co-segregated markers Unig24C11 and Gafb28I12, about 6.5 cM away; and on the other side by co-segregating markers Gafb29C08 and pAR10F10, about 9.0 cM away. *Fbl* was flanked on one side by the same marker (Gafb29C08) as N_1, about 5 cM away, but on the other side by a marker which did not show polymorphism in N_1, pAR0244 about 1.9 cM away. This region showed very high marker density in the reference map, and was tentatively inferred to be centromeric (Rong et al. 2004).

SMA-4 *SMA-4* mapping used the F_2 interspecific *G. arboreum* by *G. herbaceum* population, and a comprehensive linkage map consisting of 275 loci. The naked

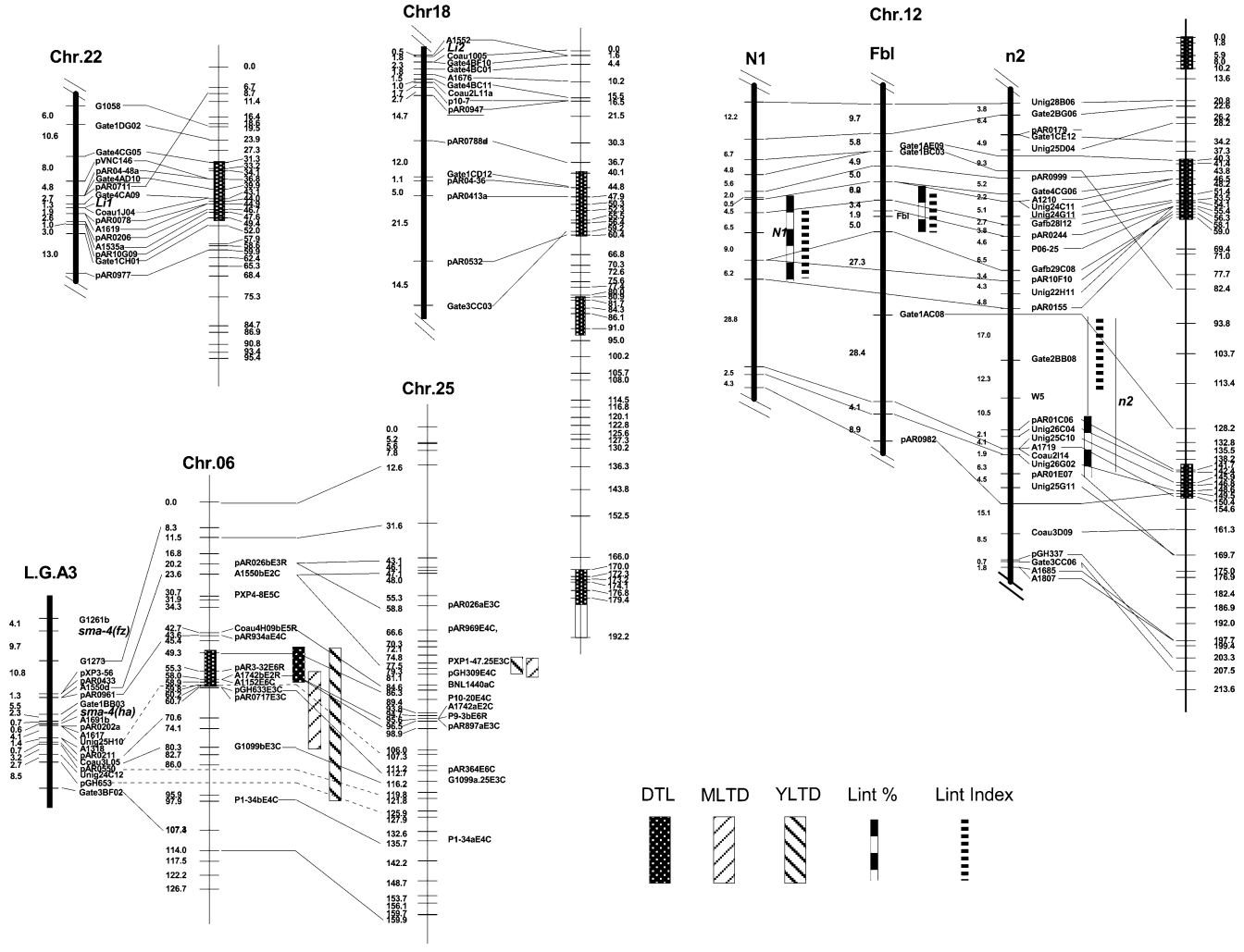

Fig. 2 Genetic mapping of cotton fiber mutants. The markers underlined are those first showing linkage between fiber mutants and RFLP markers in a genome-wide survey. Maps with thin lines are from the published reference map (Rong et al. 2004). Common loci between the different maps were connected with lines. The regions with hatched or open boxes on the reference maps contain significant ($P < 0.01$) excess or deficiency, respectively, of marker loci. The markers listed in the published Chr. 6 and Chr. 25 maps were those used to map genes affecting trichomes (Wright et al. 1999), except for BNL1440 which was the closest marker flanking *SS3* (Han et al. 2004). Coau4H09 on Chr. 6 was also mapped to the same location as BNL1440 on Chr. 25. The hatched boxes beside the Chr. 6 and Chr. 25 maps are the previously-published locations of QTLs related to the distribution of trichomes (Wright et al. 1999). The legends for QTLs of Lint % and Lint Index were presented in the figure. For n_2, an additional whisker is shown to represented LOD = 2.5–3.0, as discussed in the text

phenotype, *sma-4(fz)* mapped near the terminus of L.G. A3, flanked by G1261b (4.1 cM) and G1273 (9.7 cM: Fig. 2). Among a total of 19 loci on this linkage group, 9 are shared with tetraploid homoeologous chromosomes 6 and 25. Based on this, these three chromosomes were deduced to be homoeologous. The map of L.G.A3 covered nearly all of Chr. 25 and lacked only the distal 20 cM of Chr. 06. *sma-4(fz)* is located in a region where probes were distributed sparsely.

The *sma-4* fiberless phenotype (h_a locus) mapped to the middle region of L.G.A3, flanked by Gate1BB03 and A1691b at 2.3 and 0.7 cM, respectively. The region of *sma-4*(h_a) corresponded to a high marker density region on Chr. 6, which was deduced to be the centromeric region (Rong et al. 2004) and where a gene (*t1*) conferring the distribution of trichomes on the leaves and leaf veins (DTL) had previously been mapped (Wright et al. 1999).

Quantitative variation associated with the fiber mutants

There was considerable quantitative variation in lint fiber production superimposed on the discrete effects of the mutants. While this might be expected for interspecific crosses, we were curious whether any of the mutations may have multiple effects. Lint % and Lint Index were measured (Table 3 and Fig. 3) from those seeds with lint. Li_1, Li_2 and n_2 populations have similar average lint fraction and lint weight per seed, of about 29% and 6 grams, respectively. The range of Lint % also is similar for these three populations, with standard deviations from 5.4–6.0%. The n_2 population shows

Table 3 Quantitative summary of cotton Lint % and Lint Index

Population	Lint %		Lint Index (gram)	
	Mean	SD	Mean	SD
Pima s-7× Li_1	29.1	6.0	6.1	1.2
Pima s-7× Li_2	29.8	5.6	6.5	1.3
Pima s-7× N_1	26.6	10.3	4.9	5.5
Pima s-7× *Fbl*	14.6	10.9	3.6	2.8
Pima s-7× n_2	29.0	5.4	6.2	4.0

more variation in Lint Index (SD = 4.0 gram) than Li_1 or Li_2 (1.2, 1.3, respectively). N_1 has lower Lint % and Lint Index than Li_1, Li_2 and n_2, at about 26.6% and 4.9 gram, and with about twice as much variation in each trait. *Fbl* has by far the lowest Lint % (14.6%), only about one-half that of the other populations, but with variation similar to N_1 (SD = 10.9%). *Fbl* Lint Index is also the lowest although its variation in this trait is intermediate among the other populations.

Quantitative variation in Lint % and Lint Index is illustrated in Fig. 3, using the *Fbl* population as an example. All wild type plants are at the upper end of the range of values for Lint % and Lint Index. This figure also illustrates the general correlation between Lint % and Lint Index, which was significant for the *Fbl*, Li_1, and Li_2 populations and ranged from r = 0.873 (*Fbl*) to 0.240 (Li_2).

QTL analysis indicated that Lint % and Lint Index were closely related to the N_1, *Fb* and n_2 mutants. A major QTL, which explains 33.6% of variation of Lint % (LOD = 7.50), was mapped in the region where N_1 was located. A QTL for Lint Index was also mapped in this region [Percent variance explained (PVE) = 11.3, LOD = 3.17]. Two stronger QTLs for Lint % (PVE = 41.5, LOD = 12.68) and Lint Index (PVE = 73.2, LOD = 10.25) were detected in the same region in the *Fbl* population.

QTL analysis was especially important in the n_2 population, in which we suggest (above) incomplete expressivity and/or imperfect penetrance. A QTL explaining 20.1% of variation in Lint Index (LOD = 4.63) covers the region from pAR0155 to W5, including the possible location of n_2. A marginally-significant QTL for Lint % (PVE = 12.9%; LOD = 3.44) was also found in a region of 18.9 cM from pAR01C06 to pAR01E07 on Chr. 12. If LOD = 2.5 was used as threshold instead of 3.0, the QTL location extended upstream to pAR0155, including the possible location of n_2 (Fig. 2).

Discussion

Chromosomal locations of the fiber mutants

In this study, seven mutant genes were mapped on the cotton chromosomes, and six have been delineated by flanking DNA markers that are about 10 cM away or less. All chromosomal identities were consistent with prior chromosomal assignments (Endrizzi et al. 1985; Kohel et al. 2002), with the curious exception of n_2. Although n_2 was previously assigned to Chr. 26 (Endrizzi and Ramsay 1980), we mapped it to the homoeolog of Chr. 26, Chr. 12. Our n_2 mutant stock shows the correct phenotype, and our identification of Chr. 12 is supported by many other data (Rong et al. 2004) including independent verification (Lacape et al. 2003). We have considered the possibility that the unusual Pima S7 phenotype may somehow influence this result. If our parental stock actually were mutant at a Chr. 26 locus, and Pima S7 were also mutant at the same locus, one would expect no phenotypic segregation at that locus as we observed. However, this still fails to explain the observed segregation at the Chr. 12 locus. These observations lead us to suspect that either aneuploid analysis (Endrizzi et al. 1985) may have been erroneous, or that the source genetic stock did not actually contain

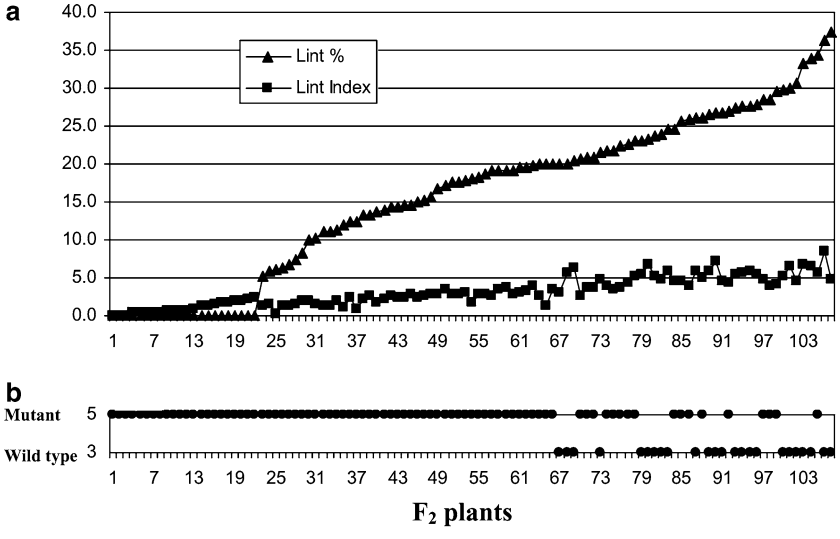

Fig. 3a–b Lint yield of F_2 plants from cross between Pima S7 and *Fbl* mutant. a Lint % and Lint Index (gram) for wild and mutant *Fbl* F_2 cotton plants. b Genotype of different F_2 plants

n_2. If the latter is true, then we have reported the location and effects of a new mutant in fiber development, although from an unknown source.

Curiously, three of the 5 tetraploid mutants map to the same chromosome (12). Two (N_1, *Fbl*) show similar phenotypes and map to locations that are indistinguishable based on present data. Although their qualitative phenotypes are similar, the demonstration that they show different quantitative phenotypes (especially Lint %) seems to confirm that they are at least different alleles, if not different genes. The observation that they fall in a chromosomal region in which genes tend to be clustered, possibly due to restricted recombination in association with the presence of a centromere, may cause close linkage between genes that are physically separated by considerable distances. We are not aware of prior allelism tests between these mutants, but plan them in view of this finding. The third mutant, n_2, a recessive allele conferring naked seed (similar to N_1 but with no loss of lint fiber: Fig. 1), shows variable expressivity but maps well outside the N_1 /*Fbl* region. If our stock truly contains n_2 (see above), then this seems to rule out the possibility that N_1 and n_2 are homoeologous, which had been suspected based on prior assignment of n_2 to Chr. 26. However, it raises the possibility that N_1 and n_2 may be related as a result of other forms of gene duplication with a proximal bias. This possibility may also apply to the *sma-4*(h_a), and *sma-4(fz)* mutants, that also map to the same chromosome as one another.

Association of fuzzless mutants with quantitative variation in lint fiber production

In addition to conferring the absence of fuzz fibers, we found the N_1, *Fbl* and n_2 mutants to map near the centers of QTL likelihood intervals for lint fiber production in the study populations (as detailed in Results). The most plausible explanation of this observation is that the same lesion that qualitatively influences fuzz fibers, also influences lint fibers but in a quantitative manner, perhaps as a result of an early perturbation in a well-buffered pathway. However, we cannot rule out the alternative that another gene in the region(s) may be involved, particularly in that the region appears to be centromeric and therefore may have a large quantity of DNA per unit of recombination. It seems unlikely that multiple independent mutations in closely-spaced genes would be implicated in the same phenotype, but cannot yet be ruled out. It is plausible that the 'wild type' parent crossed to all three mutant stocks may contribute a variant allele at a nearby locus that causes a QTL. However, other crosses between these species have not shown fiber QTLs in this region (see below). Further, it would be somewhat unexpected that a *G. barbadense* allele would increase fiber productivity relative to the corresponding *G. hirsutum* allele. While there was considerable variation in lint fiber parameters of Li_1 and Li_2, we found no QTLs corresponding to these mutants. This association thus represents another common feature of the three chromosome 12 mutants, lending some additional support to the possibility of a relationship among these genes.

Relationship of the mutants to QTLs mapped in non-mutant crosses

Prior mapping in crosses between elite (or at least advanced) forms of *G. hirsutum* and *G. barbadense* has revealed 191 fiber-related QTLs to date (Jiang et al. 1998; Saranga et al. 2001; Paterson et al. 2003; Chee et al. 2005a; Chee et al. 2005b; Draye et al. 2005). Curiously, in only one case does a discrete mutant map within the likelihood interval for a QTL discovered in a non-mutant cross, specifically Li_1 and *LgSDCT-D07* [L.G.D07 was re-assigned as Chr. 22 in the reference map (Rong et al. 2004)] which influences seed cotton yield (Jiang et al. 1998). The rarity of association between mutants and QTLs seems consistent with the hypothesis (Jiang et al 1998) that favorable alleles at genes which played major roles in the early evolution of the cotton fiber, after the divergence of the A- and F-genomes, were quickly fixed by natural selection (since spinnable fiber was established in the A-genome lineage prior to polyploid formation 1–2 million years ago, this cannot have involved human selection). Further, new alleles at these loci may perhaps be quickly eliminated from the gene pool by either natural or human selection, explaining the relative rarity of QTLs at these loci (Jiang et al. 1998). Thus, while identification of these genes is of central importance to understanding the evolution and development of the cotton fiber, it remains an open question whether they contribute appreciably to naturally-occurring variation among elite or exotic cottons. On a related note, we tacitly assume that the observation that four of the five tetraploid mutants mapped to date are genetically dominant reflects the fact that they were found in largely-homozygous cultivars, where recessive mutants would require additional generations of selfing to be detected, if they survived.

Approaches to isolate the mutated fiber genes

The present mapping of seven fiber mutants provides a starting point for fine-scale mapping and, if necessary, positional cloning. The Li_2 mutation is especially promising, with markers 0.7 and 1.8 cM away (as estimated based on 158 F_2 plants) and located near the end of its chromosome (18) where the relationship between genetic and physical distance should be favorable. While we hope that other avenues of study will provide complementary data that obviate the need to 'walk' through the centromeric regions which contain N_1, *Fbl*, and Li_1, if necessary large numbers of recombinants can be obtained.

The cotton fiber transcriptome contains an estimated 36,000 homoeologous gene transcripts from the At and

Dt subgenomes in tetraploid cotton (Arpat et al. 2004; Wilkins and Arpat 2005). Identifying candidate genes that play a pivotal role in fiber development, and in the determination of important agronomic fiber traits poses a challenge to cotton researchers. Research into the genetic control of cotton fiber development may therefore benefit greatly from progress in understanding the growth and development of hair-bearing epidermal cells (trichomes) in *Arabidopsis*, thought to have shared common ancestor with *Gossypium* about 85 million years ago (Benton 1993). Among at least 27 genes known to regulate aspects of *Arabidopsis* trichome development (Larkin et al. 2003; Schiefelbein 2003), many are orthologs or close homologs to cotton fiber genes (Arpat et al. 2004) suggesting some general parallels in control of trichome morphogenesis in *Arabidopsis* and cotton, although a number of regulatory genes have been identified that are novel and specific to fiber and/or cotton (Loguercio et al. 1999; Wilkins and Jernstedt 1999; Arpat et al. 2004). Conservation of synteny between *Arabidopsis* and cotton (Bowers et al. 2003) is likely to become clearer as more data is acquired about an ancient duplication specific to the cotton lineage (Rong et al. 2005). This may help to accelerate identification of candidate genes (Paterson et al. 1995; Paterson et al. 1996) or molecular markers in a region of interest (Oh et al. 2002).

A growing body of genomic data for cotton itself provides additional avenues to explore for candidate genes that may account for mutations in fiber development. For example, an EST (NAU1272) recently mapped on Chr. 6 (Han et al. 2004) was annotated as Sucrose synthase (*Sus*), which is thought to play a crucial role in fiber cell initiation and elongation (Ruan et al. 2003). In our high-density map (Rong et al. 2004), BNL1440, the closest marker to *Sus* on Chr. 6, was mapped on homoeologous Chr. 25 (Fig. 2). A marker (Coau4H09), co-segregated with BNL1440 on Chr. 25, and detected a homoeologous site on Chr. 6 about 0.9 cM above pAR0934 which was diagnostic of the *t1* trichome gene (Wright et al. 1999). Interestingly, two QTLs related to trichome density on the young and mature leaves, inferred to be the *t2* locus, also mapped very close to BNL1440 on Chr. 25 (Wright et al. 1999), in the region homoeologous to *Sus*. One qualitative (distribution of trichomes on leaves, DTL, inferred to be t_1) and two quantitative phenotypes for trichome distribution mapped in the homoeologous region, below pAR0934 on Chr. 6 (Wright et al. 1999). Finally, this location corresponds very closely to the A-genome location of the *sma-4(fbl)* trait mapped herein. Several prior reports (Simpson 1947; Knight 1952; Kloth 1995) also suggest association of the t_1 trichome locus to fiber quality, an association that we extend herein to include *sma-4(fbl)*, *Sus*, and perhaps t_2. This close positional association between sma-$4(h_a)$, t_1, t_2, and sucrose synthase, raises the question of whether these genes may be functionally related. Moreover, this association illustrates the principle that the convergence of a growing number of diverse data types promises new hypotheses leading to accelerated genetic dissection of cotton fiber development.

Acknowledgements We thank members of the Paterson lab for many valuable contributions, and the US National Science Foundation (J.F.W., T.A.W., A.H.P.), USDA National Research Initiative (J.F.W., A.H.P.), and BOYSCAST program of the DST, India (V.N.W.) for financial support.

References

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

Arpat A, Waugh M, Sullivan JP, Gonzales M, Frisch D, Main D, Wood T, Leslie A, Wing R, Wilkins T (2004) Functional genomics of cell elongation in developing cotton fibers. Plant Mol Bio 54:911–929

Basra AS, Malik CP (1984) Development of the cotton fiber. Int Rev Cytol 89:65–113

Beasley CA, Egli E (1977) Fiber production in vitro from a conditional fiberless mutant of cotton. Dev Biol 57:234–237

Benton MJ (1993) The fossil record 2. Chapman and Hall, New York

Bowers JE, Chapman BA, Rong JK, Paterson AH (2003) Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. Nature 422:433–438

Chee P, Draye X, Jiang C, Decanini L, Delmonte T, Bredhauer B, Smith CW, Paterson H (2005a) Molecular dissection of interspecific variation between Gossypium hirsutum and G. barbadense (cotton) by a backcross-self approach: I. Fiber elongation. Theor Appl Genet (in press)

Chee P, Draye X, Jiang C, Decanini L, Delmonte T, Bredhauer B, Smith CW, Paterson H (2005b) Molecular dissection of interspecific variation between Gossypium hirsutum and G. barbadense (cotton) by a backcross-self approach: III. Fiber length. Theor Appl Genet (in press)

Cronn RC, Small RL, Haselkorn T, Wendel JF (2002) Rapid diversification of the cotton genus (Gossypium : Malvaceae) revealed by analysis of sixteen nuclear and chloroplast genes. Am J Bot 89:707–725

Draye X, Chee P, Jiang C, Decanini L, Delmonte T, Bredhauer B, Smith CW, Paterson H (2005) Molecular dissection of interspecific variation between Gossypium hirsutum and G. barbadense (cotton) by a backcross-self approach: II. Fiber fineness. Conditionally accepted, pending revision. Theor Appl Genet (in press)

Endrizzi JE, Ramsay G (1980) Identification of ten chromosome deficiencies in cotton. J Hered 71:45–48

Endrizzi JE, Turcotte EL, Kohel RJ (1985) Genetics, cytogenetics, and evolution of Gossypium. Adv Genet 23

Fryxell PA (1963) Morphology of the base of seed hairs of Gossypium. Bot Gaz 124:196–199

Fryxell P (1979) The Natural History of the Cotton Tribe. Texas A&M University Press, College Station, TX

Griffee F, Ligon LL (1929) Occurrence of lintless cotton plants and the inheritance of the character 'lintless.' J Amer Soc Agron 21:711–717

Han ZG, Guo WZ, Song XL, Zhang TZ (2004) Genetic mapping of EST-derived microsatellites from the diploid Gossypium arboreum in allotetraploid cotton. Mol Gen Genet 272:308–327

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

Karaca M, Saha S, Jenkins JN, Zipf A, Kohel R, Stelly DM (2002) Simple Sequence Repeat (SSR) Markers Linked to the Ligon Lintless (Li1) Mutant in Cotton. J Hered 93:221–224

Kearney TH, Harrison GJ (1927) Inheritance of smooth seeds in cotton. J Agr Res 35:193–217

Kloth R (1995) Quantitative trait loci affecting cotton fiber are linked to the T(1) locus in upland cotton. Theor Appl Genet 91:762–768

Knight RL (1952) The genetics of jassid resistance in cotton. I. The genes H_1 and H_2. J Genet 51:46–66

Kohel RJ (1979) Gene arrangement in the duplicate linkage group V anf IX: Nectariless, glandless, and withering bract in cotton (Gossypium hirsutum L.). Crop Sci 19:831–833

Kohel RJ, Narbuth EV, Benedict CR (1992) Fiber development of Ligon lintless-2 mutant of cotton. Crop Sci 32:733–735

Kohel RJ, Stelly DM, Yu J (2002) Tests of six cotton (Gossypium hirsutum L.) mutants for association with aneuploids. J Hered 93:130–132

Kosambi D (1944) The estimation of map distance from recombination values. Ann Eugen 12:172–175

Lacape JM, Nguyen TB, Thibivilliers S, Bojinov B, Courtois B, Cantrell RG, Burr B, Hau B (2003) A combined RFLP-SSR-AFLP map of tetraploid cotton based on a Gossypium hirsutum x Gossypium barbadense backcross population. Genome 46:612–626

Lander ES, Botstein D (1989) Mapping Mendelian Factors Underlying Quantitative Traits Using Rflp Linkage Maps. Genetics 121:185–199

Lander E, Green P, Abrahamson J, Barlow A, Daly M, Lincoln S, Newburg L (1987) MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181

Larkin JC, Brown ML, Schiefelbein J (2003) How do cells know what they want to be when they grow up? Lessons from Epidermal Patterning in Arabidopsis. Annual Review of Plant Biology 54:403–430

Loguercio LL, Scott HC, Trolinder NL, Wilkins TA (1999) Hmg-co-A reductase gene family in cotton (*Gossypium hirsutum* L.): Unique structural features and differential expression of hmg2 are potentially associated with synthesis of specific isoprenoids in developing embryos. Plant Cell Physiol 40:750–761

Moulherat C, Tengberg M, Haquet JF, Mille B (2002) First evidence of cotton at Neolithic Mehrgarh, Pakistan: Analysis of mineralized fibres from a copper bead. J Archaeol Sci 29:1393–1401

Narbuth EV, Kohel RJ (1990) Interitance and linkage analysis of a new fiber mutant in cotton. J Hered 81:131–133

Oh KC, Hardeman K, Ivanchenko MG, Ellard-Ivey M, Nebenführ A, TJ White TJ, Lomax TL (2002) Fine mapping in tomato using microsynteny with the Arabidopsis genome: the Diageotropica (Dgt) locus. Genome Biol 3: research0049.0041–0049.0011

Paterson A, Lin Y-R, Li Z (1995) Convergent domestication of cereal crops by independent mutations at corresponding genetic loci. Science 269:1714–1718

Paterson AH, Lan TH, Reischmann KP, Chang C, Lin YR, Liu SC, Burow MD, Kowalski SP, Katsar CS, DelMonte TA, Feldmann KA, Schertz KF, Wendel JF (1996) Toward a unified genetic map of higher plants, transcending the monocot-dicot divergence. Nat Genet 14:380–382

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

Percy RG, Kohel RJ (1999) Qualitative Genetics. In: Smith CW, Cothren JT (eds) Cotton: Origin, History, Technology, and Production. John Wiley & Sons, pp 319–360

Reinisch A, Dong J-M, Brubaker C, Stelly D, Wendel J, Paterson A (1994) A detailed RFLP map of cotton (Gossypium hirsutum x G. barbadense): Chromosome organization and evolution in a disomic polyploid genome. Genetics 138:829–847

Rong J-K, Abbey C, Bowers JE, Brubaker CL, Chang C, Chee PW, Delmonte TA, Ding XL, Garza JJ, Marler BS, Park C-H, Pierce GJ, Rainey KM, Rastogi VK, Schulze SR, Trolinder NL, Wendel JF, Wilkins TA, Williams-Coplin TD, Wing RA, Wright RJ, Zhao X, Zhu L, Paterson AH (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

Rong J-K, Bowers JE, Schulze SR, Waghmare VN, Rogers CJ, Pierce GJ, Zhang H, Estill JC, Paterson AH (2005) Comparative Genomics of Gossypium and Arabidopsis: Unraveling the Consequences of both Ancient and Recent Polyploidy. Genome Res (in press)

Ruan Y-L, Gilmore L, Conner T (1998) Towards *Arabidopsis* genome analysis: monitoring expression profiles of 1400 genes using cDNA microarrays. Plant J 15:821–833

Ruan Y-L, Llewellyn DJ, Furbank RT (2003) Suppression of Sucrose Synthase Gene Expression Represses Cotton Fiber Cell Initiation, Elongation, and Seed Development. Plant Cell 15:952–964

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

Schiefelbein J (2003) Cell-fate specification in the epidermis: a common patterning mechanism in the root and shoot. Current Opinion in Plant Biology 6:74–78

Silow RA (1941) The comparative genetics of Gossypium anomalum and the cultivated Asiatic cottons. J Genet 42:259–358

Simpson DM (1947) Fuzzy leaf in cotton and its association with short lint. J Hered 38:153–156

Turley RB, Ferguson DL (1996) Changes of ovule proteins during early fiber development in a normal and a fiberless line of cotton (Gossypium hirsutum L). J Plant Physiol 149:695–702

Wang XD, Zhu YX, Ji DF, Jiang SL, Li YY (2001) Cloning of fiber-specific cDNAs and their structural variations in 4 fiber mutants. Chinese Science Bulletin 46:234–237

Wendel JF (1989) New World Tetraploid Cottons Contain Old-World Cytoplasm. Proc Natl Acad Sci USA 86:4132–4136

Wilkins TA, Arpat A (2005) The cotton fiber transcriptome. Physiol Plantarum: in press

Wilkins TA, Jernstedt JA (1999) Molecular genetics of developing cotton fibers. In: Basra AS (eds) Cotton Fibers. Haworth Press, New York, pp 231–267

Wilkins TA, Arpat A, Sickler BA (2005) Cotton fiber genomics: developmental mechanisms. Pflanzenschutz-Nachricten 58:119–139

Wright RJ, Thaxton PM, El-Zik KH, Paterson AH (1999) Molecular mapping of genes affecting pubescence of cotton. J Hered 90:215–219

Zhang T, Yuan Y, Yu J, Guo W, Kohel RJ (2003) Molecular tagging of a major QTL for fiber strength in Upland cotton and its marker-assisted selection. Theor Appl Genet 106:262–268