Check for updates

The *Ligon lintless*-2 Short Fiber Mutation Is Located within a Terminal Deletion of Chromosome 18 in Cotton[1]

Jinesh D. Patel,[a,2] Xianzhong Huang,[a,b] Lifeng Lin,[a] Sayan Das,[a] Rahul Chandnani,[a] Sameer Khanal,[a] Jeevan Adhikari,[a] Tariq Shehzad,[a,d] Hui Guo,[a] Eileen M Roy-Zokan,[a] Junkang Rong,[c] and Andrew H. Paterson[a,3,4]

[a]Plant Genome Mapping Laboratory, University of Georgia, Athens, Georgia 30602
[b]Plant Genomics Laboratory, College of Life Sciences, Shihezi University, 832003 Shihezi, China
[c]Zhejiang A&F University, Linan, Hangzhou 311300, Zhejiang, China
[d]Department of Biological and Environmental Sciences, College of Arts and Sciences, Qatar University, 2713 Doha, Qatar

ORCID IDs: 0000-0001-6696-2155 (J.D.P.); 0000-0003-1264-8106 (X.H.); 0000-0002-7981-7296 (S.D.); 0000-0002-3342-0381 (S.K.); 0000-0002-3815-3067 (J.A.); 0000-0003-2159-0487 (A.H.P.).

Extreme elongation distinguishes about one-fourth of cotton (*Gossypium* sp.) seed epidermal cells as "lint" fibers, useful for the textile industry, from "fuzz" fibers (<5 mm). *Ligon lintless-2* (Li_2), a dominant mutation that results in no lint fiber but normal fuzz fiber, offers insight into pathways and mechanisms that differentiate spinnable cotton from its progenitors. A genetic map developed using 1,545 F2 plants showed that marker CISP15 was 0.4 cM from Li_2, and "dominant" simple sequence repeat (SSR) markers (i.e. with null alleles in the Li_2 genotype) SSR7 and SSR18 showed complete linkage with Li_2. Nonrandom distribution of markers with null alleles suggests that the Li_2 phenotype results from a 176- to 221-kb deletion of the terminal region of chromosome 18 that may have been masked in prior pooled-sample mapping strategies. The deletion includes 10 genes with putative roles in fiber development. Two Glycosyltransferase Family 1 genes showed striking expression differences during elongation of wild-type versus Li_2 fiber, and virus-induced silencing of these genes in the wild type induced Li_2-like phenotypes. Further, at least 7 of the 10 putative fiber development genes in the deletion region showed higher expression in the wild type than in Li_2 mutants during fiber development stages, suggesting coordinated regulation of processes in cell wall development and cell elongation, consistent with the hypothesis that some fiber-related quantitative trait loci comprise closely spaced groups of functionally diverse but coordinately regulated genes.

The ability of single cotton (*Gossypium* sp.) fiber cells to reach as much as 6 cm in length confers their value to the textile industry (Kim and Triplett 2001). Cotton fiber cell development is widely conceptualized as comprising four stages, i.e. initiation, elongation, secondary wall biosynthesis, and maturation (Basra and Malik 1984). Fiber elongation typically lasts for 15 to 20 d, during which the fiber expands at a rate of 2 mm per day (Lee et al., 2007). The process involves metabolic pathways regulating hydrogen peroxide (H_2O_2) and reactive oxygen species, Ca^{2+}, stress, brassinosteroids, ethylene, water transportation, cell wall loosening, and pectin biosynthesis (Hovav et al., 2008; Pang et al., 2010; Li et al., 2013; Fang et al., 2014; Shan et al., 2014; Tang et al., 2014; Yang et al., 2014). Secondary cell wall synthesis overlaps with fiber elongation, during which large amounts of cellulose are synthesized and deposited, ultimately leading to thickening of the cell wall to ~3 to 4 μm (Kim and Triplett, 2001). Cell wall thickening is necessary to impart fiber strength, which is an important component of cotton fiber quality (Pang et al., 2010). The final step of fiber development is fiber maturation, which may last until 60 d post-anthesis (DPA). At the end of the process, ~500,000 elongated lint fibers will have been produced in a single fruit ("boll") containing 30 to 35 seeds (Bowman et al., 2001).

[1]This work was financially supported by a University System of Georgia Regents Professorship to A.H.P., the Georgia Agricultural Experiment Station (to A.H.P.), and a postdoctoral fellowship from the China Scholarship Council (to X.H.).

[2]Present address: Department of Crop, Soil and Environmental Sciences, Auburn University, Auburn, AL 36849.
[3]Author for contact: paterson@uga.edu.
[4]Senior author.
The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: Andrew H. Paterson (paterson@uga.edu).

J.D.P. designed, performed the experiments, carried out data analysis and wrote the manuscript; X.H. assisted with designing, performing, and analyzing data for reverse-transcription quantitative PCR and virus-induced gene silencing experiments; L.L. performed initial screening and fine mapping; S.D., R.C., S.K., and J.A. helped with field work and marker screening; T.S. assisted in reverse-transcription quantitative PCR and virus-induced gene silencing experiments; H.G. helped with SNP calling between fiber and nonfiber groups for candidate genes; E.M.R. helped with phylogenetic analysis; R.K. helped with population development and initial genetic mapping; and A.H.P. conceived the project, participated in the experimental design and supervised the experiments, oversaw the data analysis, and revised the manuscript before submission.

www.plantphysiol.org/cgi/doi/10.1104/pp.19.01531

 © 2020 American Society of Plant Biologists. All Rights Reserved.

Downloaded from https://academic.oup.com/plphys/article/183/1/277/6116371 by Qatar University user on 14 March 2024

Cotton mutants with fiber anomalies are excellent tools to use in deciphering the complex process of fiber development. Three mutants with fuzzy seeds and short lint fiber have been identified in cotton, namely, *Ligon lintless-1* (Li_1; Griffee and Ligon 1929), Li_2 (Narbuth and Kohel 1990), and the *Ligon lintless-like* mutant (Li_x; Cai et al., 2013), mapping to chromosomes 22 (D4), 18 (D13), and 4 (A4), respectively (Rong et al., 2005; Cai et al., 2013). While the Li_1 and Li_x mutants have deformed leaves and stem, Li_2 plants show normal vegetative growth. Multiple studies have revealed the genetic position, differential gene expression, and metabolite changes in this mutant, but have not yet identified the mutation responsible for the Li_2 phenotype (Hinchliffe et al., 2011; Gilbert et al., 2013; Naoumkina et al., 2013; Thyssen et al., 2014).

We sought to identify the Li_2 gene by integrating positional (genetic and physical mapping) information, evolutionary information, differential gene expression, and functional validation using virus-induced gene silencing (VIGS). Identification of seven genes in the target region that are coordinately expressed in fiber development supports a recent model (Paterson et al., 2012) suggesting that many cotton fiber-related quantitative trait loci (QTLs) may be complex. In addition, VIGS reduced fiber length in two genes, suggesting the identity of Li_2 and the functions of genes related to elongation of one of the longest known single cells in the plant kingdom.

RESULTS

Phenotyping of Li_2 Mutants and Segregation of the Li_2 Region

The Li_2 mutant has seeds with no lint fiber but normal fuzz fiber. To determine whether a plant was Li_2 or wild type, multiple open bolls were inspected on each plant (Fig. 1). Some plants showed both Li_2 and wild-type bolls (Fig. 1), in which case the plant was considered Li_2. Much literature has shown the Li_2 mutant phenotype to be genetically dominant and governed by a single locus; thus, a segregation ratio of three lintless to one linted individual(s) was expected (Narbuth and Kohel 1990; Rong et al., 2005). However, among 1,545 F2 plants we found 1,091 to be lintless and 454 to be wild type, a 2.4:1 ratio, suggesting enrichment of the wild type relative to the Mendelian expectation. As shown in Table 1, a similar pattern was also observed for DNA marker genotypes in the Li_2 region. Such segregation distortion may be due to the presence of favorable alleles in this region from *Gossypium barbadense*, or to a deleterious effect of the Li_2 mutant (from *Gossypium hirsutum*).

Chromosome Walking and Fine Mapping of the Li_2 Locus

Initial mapping of the Li_2 locus in 135 individuals showed 12 markers to be toward the centromeric side of the gene, i.e. lacking flanking markers on the telomeric side that would permit identification of recombinants in the interval containing the gene. To try to develop markers on the telomeric side of Li_2, we designed overgo probes using genetic sequences of markers closest to the gene (Gr_ea17f11, A1552, Gh.fbr.sw02661, and COAU2K07), hybridized them to *Gossypium raimondii* bacterial artificial chromosomes (BACs; Supplemental Table S1), and sequenced BACs GR174F21, GR109E22, and GR174O23 (Supplemental Fig. S1A). Based on alignment and segregation of markers developed from these BACs with Li_2, GR109E22 and GR174O23 overlapped each other, but there was a gap between

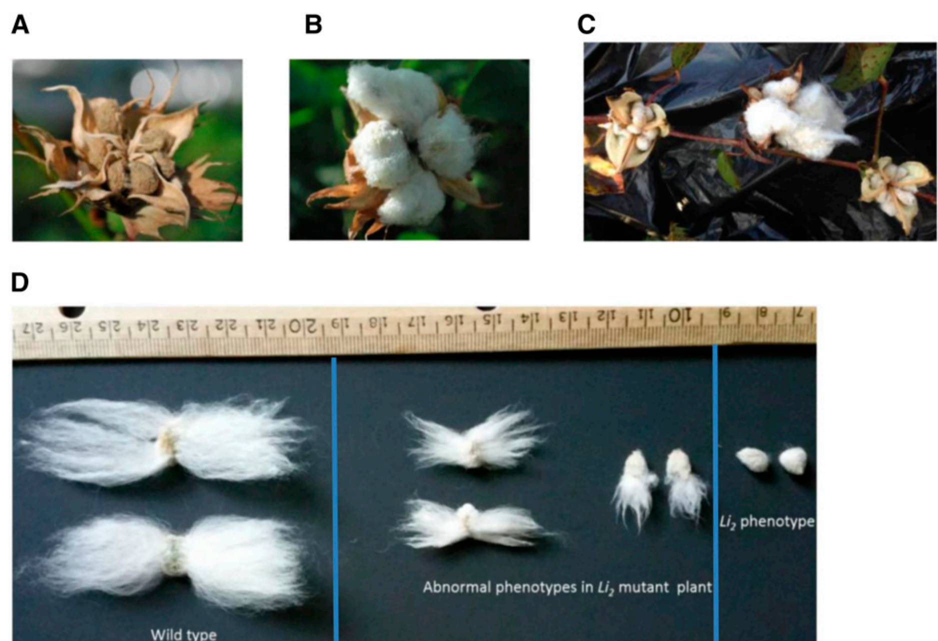

Figure 1. Phenotype of seeds containing the Li_2 mutant allele, and homozygous wild-type allele. A, Li_2 mutant allele. B, Homozygous wild-type allele. C, A branch showing atypical expression of the Li_2 phenotype (two bolls of Li_2 phenotype on either side of the normal lint or wild-type phenotype). D, Atypical expression of Li_2 phenotype within a boll. The left and right show the wild-type and Li_2 homozygotes, respectively, while the center shows the two abnormal phenotypes seen in some bolls of Li_2 plants.

Downloaded from https://academic.oup.com/plphys/article/183/1/277/6116371 by Qatar University user on 14 March 2024

Table 1. *Segregation distortion of DNA markers in the* Li_2 *region*

Genotype	CISP15			NAU2980			NAU3827		
	Observed	Expected	Chi test	Observed	Expected	Chi test	Observed	Expected	Chi test
GH/GH	284	380.00	2.02E-08	285	382.75	2.16E-08	288	383.50	7.56E-08
GH/GB	796	760.00	–	808	765.50	–	814	767.00	–
GB/GB	440	380.00	–	438	382.75	–	432	383.50	–
Total	1,520	–	–	1,531	–	–	1,534	–	–

GR109E22 and GR174F21, with the latter being closest to Li_2. When we BLASTed 13,662 BAC end sequences from the *G. raimondii* library (Lin et al., 2010) to seek BACs with one end in the terminal region of chromosome D13, we found GR102N11 and GR006L12, but DNA fingerprinting showed that sequencing these BACs would not improve coverage of the region (Supplemental Fig. S1B). GR174F21 was later found to extend from 60,413,412 to 60,534,221 bp of chromosome D13 of *G. hirsutum*, with just 78 bp beyond it to the end of the chromosome (Zhang et al., 2015). Thus, GR174F21 was the last informative BAC sequence for the terminal region of chromosome D13.

For fine mapping, four markers, NAU2980, NAU3447, NAU3827, and NAU3223, covering the Li_2 region were used along with new simple sequence repeat (SSR) markers developed from BAC GR109E22 (CISP15) and GR174F21 (SSR7 and SSR18). A total of 1,545 F2 plants were genotyped. Based on the genetic map, CISP15 was 0.4 cM from Li_2, while SSR7 and SSR18 cosegregated with Li_2 (Fig. 2). Both cosegregating markers were dominant and in repulsion phase with Li_2.

Deletion of a Chromosome Segment Appears Responsible for the Li_2 Phenotype

To try to mitigate the constraint that SSR7 and SSR18 from BAC GR174F21 were genetically dominant, two more SSRs (CISP39 and gr174F21_4.4) were found to be polymorphic and map near Li_2. However, while the majority of SSRs in other regions had been codominant, these (like SSR7 and SSR18) were also dominant, all four amplifying only *G. barbadense* alleles. This observation suggested that a segmental deletion near the chromosome 18 terminus may account for the Li_2 mutant, with its genetic dominance due to haploinsufficiency, i.e. a single functional allele does not produce enough gene product to confer a wild-type phenotype. To further investigate the hypothesis of a deletion in the Li_2 region of chromosome 18, we sequenced additional segments in this region from the wild type (*G. barbadense)* and homozygous Li_2 mutants (*G. hirsutum*). The D-genome sequence was mined for genic sequences beyond the last codominant marker, i.e. CISP15 toward the telomere. Although precautions were taken while designing primers and running PCR, only five gene primer sets amplified single bands useful for further study. One (Gorai.013G268200) was between CISP15 and the centromere; and four, Gorai.013G269200, Gorai.013G270000, Gorai.013G270800, and Gorai.013G271400, were between CISP15 and the telomere. As Li_2 lies on a chromosome of the Dt genome, the amplicon sequence in the hypothetical deletion region from Li_2 homozygous mutants should only show sequence from the At genome, while wild-type plants should show sequences from both At and Dt genomes. Sequencing of Gorai.013G268200, which is between the centromere and codominant marker CISP 15, suggested the presence of both genome (At and Dt) segments in the mutant and wild type. Furthermore, due to stringency of the primers of Gorai.013G269200, only Dt was amplified in the Li_2 and wild-type plants, suggesting no deletion in the Li_2 mutant until this point. Amplicons of three genes, Gorai.013G270000, Gorai.013G270800, and Gorai.013G271400, were missing Dt segments in Li_2 homozygous lines, with only genotypes corresponding to the At genome present in the mutants, whereas both At and Dt genotypes were present in the wild type (Supplemental Fig. S2). Thus, three consecutive primer pairs fail to amplify a chromosome 18 (Dt) locus from the *G. hirsutum* genotype carrying Li_2, while Gorai.013G268200 and Gorai.013G269200 (closer to CISP15 than the other three) amplify Dt loci, supporting the hypothesis that a terminal deletion is responsible for the Li_2 phenotype. The results suggest the deletion breakpoint to be between Gorai.013G269200 and Gorai.013G270000, with at least 176 kb and perhaps as much as 221 kb missing from the terminal end of chromosome D13 in Li_2 mutants.

Candidate Li_2 Gene Sequences

The putative 176- to 221-kb deletion of chromosome 18 in the Li_2 mutants spanned a total of 35 genes (Supplemental Table S2), among which 10 had annotations suggesting that they function as cell wall proteins and/or enzymes, or are associated with secondary metabolism, hormonal regulation, and posttranscriptional modification. To further narrow down the list we compared coding sequences and untranslated regions (5′ and 3′ UTRs) of the 10 genes between short fiber-producing (D and F) and fiber-producing (A, At, and Dt) genomes. We found 15 single-nucleotide polymorphism (SNP) alleles differentiating nonfiber and fiber-producing groups, with 11 in the coding regions of eight genes inferred to cause amino acid changes and potentially affect gene function (Table 2).

Downloaded from https://academic.oup.com/plphys/article/183/1/277/6116371 by Qatar University user on 14 March 2024

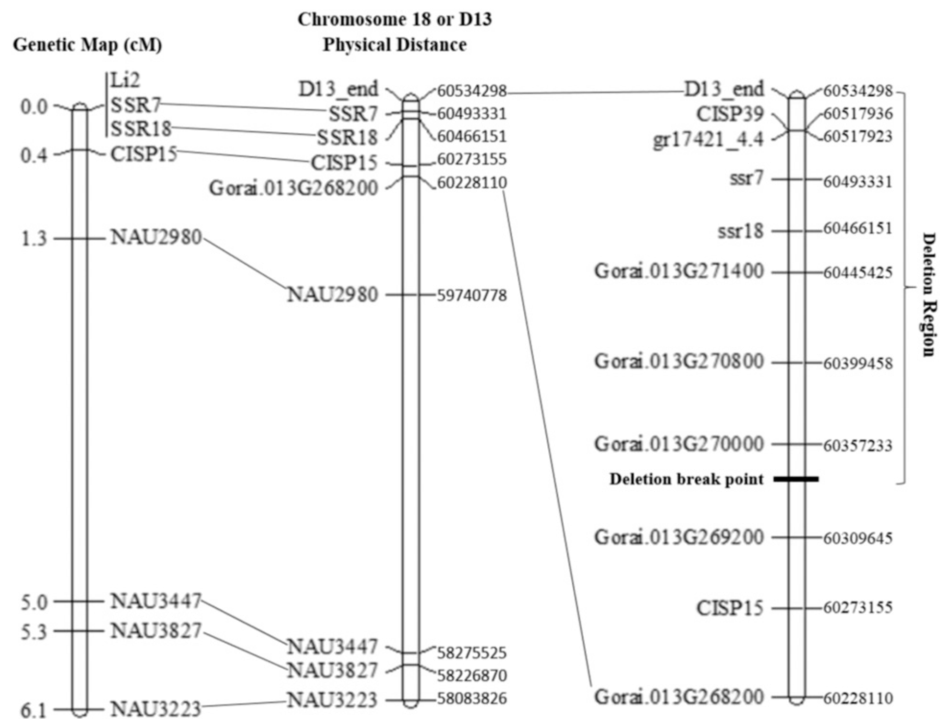

Figure 2. Genetic and physical map of the Li_2 locus. Genetic distance is in centimorgans and physical distance is in base pairs. An approximate deletion break point is shown. Markers from the deletion region (ssr7 and ssr18) are completely linked with the Li_2 phenotype. Gorai.013Gxxxxxx represents genes from which amplicons were sequenced to investigate the deletion theory and find approximate break points.

Differential Expression of the Candidate Gene during Fiber Development Based on RT-qPCR

Based on reverse-transcription quantitative PCR (RT-qPCR) at seven different fiber development stages (Fig. 3), the 10 candidate genes in the putative Li_2 deletion showed five broad categories of expression pattern. Two genes (*GhUBE11-D1a* and *GhC4H*) were both expressed primarily at 3 and 6 DPA, with substantially diminished expression later, and with only small differences between Li_2 and the wild type. Four genes (*GhUGT87A1-D1a*, *GhUGT87A2*, *GhUGT87A1-D1b*, and *GhETO1*) were expressed at low levels at 3 DPA, with much higher levels at 6 and 12 DPA (and higher in the wild type than in the Li_2 mutant), diminishing at 15 DPA with the wild type maintaining higher expression than the Li_2 mutant for *GhUGT87A1-D1a* and *GhETO1* (and noting that the transition from 6 to 10 DPA is confounded with sampling of ovule versus fiber tissue so is not directly comparable). Two genes (*GhE1310* and *GhEXPA8*) paralleled the general expression pattern of the second group, but with striking enrichment of expression in the Li_2 mutant at 15 DPA. *GhIRX7* showed enriched expression in the Li_2 mutant early in development (3–6 DPA) but not until later (21–24 DPA) in the wild type. *GhUBE11-D1b*, a ubiquitin-activating enzyme, was enriched in the wild type at all developmental stages except 10 DPA.

VIGS

We used VIGS to conduct functional complementation tests for 10 candidate genes in the putative Li_2 deletion region. Due to constraints on growth chamber space, we performed an initial VIGS experiment on eight plants per gene in the greenhouse with controlled temperature and humidity, finding substantial reduction in the fiber length of plants treated with TRV2:UGT87A1-D1a and TRV2:UGT87A2. In a follow-up study in a growth chamber, bolls collected from multiple plants treated with TRV2:UGT87A1-D1a showed 32% to 40% reduction in fiber length compared to control lines, with 15% to 25% reduction in fiber length for plants treated with TRV2:UGT87A2 (Fig. 4). Despite the use of a growth chamber with controlled environmental conditions validated in prior work (Gao et al., 2011), VIGS phenotypes varied somewhat between different plants and within the same plant, as has been reported previously in cotton (Wan et al., 2016; Andres et al., 2017). Similar problems were also seen in plants treated with the control vector TRV2:GRCLA1 silencing the cotton *CLOROPLASTOS ALTERADOS 1* gene (Supplemental Fig. S3; Gao et al., 2011).

Phylogenetic Analysis

Sequences of *UGT87A1* and *UGT87A2* genes were extracted from genome sequences of *G. raimondii* (Paterson et al., 2012), *Gossypium arboreum* (Li et al., 2014), and *G. hirsutum* (Zhang et al., 2015), and from resequencing data for *Gossypium longicalyx* (Paterson et al., 2012). Protein sequences were predicted using FGENESH and employed in phylogenetic analysis performed using MrBayes v3.2 (Kumar et al., 2016) with the Jones-Taylor-Thornton model of protein evolution (Jones et al., 1992). The analysis was run for

Downloaded from https://academic.oup.com/plphys/article/183/1/277/6116371 by Qatar University user on 14 March 2024

Table 2. *Properties of nonsynonymous SNPs differentiating fiber and short-fiber producing* Gossypium *sp.*
Period represents no change in the amino acid due to SNP between short-fiber and fiber producing *Gossypium* sp. CDS, Coding sequence.

Gene id (*G. ramondii*)	Position ID[a]	Strand	Non-fiber	Fiber	Snp Location	Short-fiber Amino Acid	Fiber Amino Acid	Gene ID (*G. hirsutum*)	Annotated Function
Gorai.013G269400	58104353	–	A	T	CDS	F	Y	Gh_D13G2434	Glucuronoxylan glucuronosyltransferase IRX7
Gorai.013G270200	58142691	+	G	A	CDS	G	R	Gh_D13G2442	Ethylene-overproduction protein 1-like
Gorai.013G270200	58143864	+	C	A	CDS	L	I	Gh_D13G2442	Ethylene-overproduction protein 1-like
Gorai.013G270200	58144018	+	G	A	CDS	G	D	Gh_D13G2442	Ethylene-overproduction protein 1-like
Gorai.013G270200	58144068	+	G	A	CDS	V	I	Gh_D13G2442	Ethylene-overproduction protein 1-like
Gorai.013G270300	58144807	–	A	G	3-UTR	.	.	Gh_D13G2443	Ubiquitin-activating enzyme E1 1-like
Gorai.013G270300	58144988	–	T	G	3-UTR	.	.	Gh_D13G2443	Ubiquitin-activating enzyme E1 1-like
Gorai.013G270400	58154218	–	T	A	CDS	Q	H	Gh_D13G2444	Ubiquitin-activating enzyme E1 1-like, transcript variant X2
Gorai.013G270800	58179089	+	A	G	CDS	E	G	Gh_D13G2448	Glucan endo-1,3-β-glucosidase 9-like
Gorai.013G270900	58181681	+	C	T	CDS	P	L	Gh_D13G2449	UDP-glycosyltransferase 87A1-like
Gorai.013G271000	58184415	+	C	G	CDS	R	G	Gh_D13G2450	UDP-glycosyltransferase 87A2-like, transcript variant X2
Gorai.013G271000	58185533	+	C	T	3-UTR	.	.	Gh_D13G2450	UDP-glycosyltransferase 87A2-like, transcript variant X2
Gorai.013G271100	58187286	+	T	A	CDS	I	N	Gh_D13G2452	UDP-glycosyltransferase 87A1-like
Gorai.013G271700	58235904	+	C	T	5-UTR	.	.	Gh_D13G2458	Transcinnamate 4-monooxygenase-like
Gorai.013G272000	58243663	+	A	G	CDS	R	G	Gh_D13G2460	Expansin-A8-like

[a]Position is based on D genome sequence (Paterson et al., 2012).

5,000,000 generations with sample frequency every 100 generations. The resulting tree was visualized in MEGA7 (Kumar et al., 2016). *GhUGT87A1_D1a* and *GhUGT87A2* appear more closely related to each other than either is to *GhUGT87A1_D1b*. Further protein sequence alignment using MUSCLE (Edgar 2004) showed truncation of *GhUGT87A2* in the At subgenome at the C-terminal end, with partial deletion (16 of 44 amino acids) of a highly conserved region known as plant secondary product glycosyltransferase (PSPG; Supplemental Fig. S4) that plays a role in transferring sugar molecules from donor to acceptor molecules (Osmani et al., 2009). Thus, deletion in PSPG can affect normal transfer of sugar molecules by *GhUGT87A2* of the At subgenome, causing partial or complete loss of function.

DISCUSSION

Investigation of the causal agent(s) of the cotton Li_2 mutant phenotype, conferring seeds with no lint fiber but normal fuzz fiber, provides an early example in support of the hypothesis that groups of closely spaced genes that are functionally diverse but coordinately regulated may be of central importance to cotton fiber development (Paterson et al., 2012). Two prior studies (Hinchliffe et al., 2011; Thyssen et al., 2014) had mapped Li_2 to different positions on chromosome 18 than ours. Our recombination data between SSR markers and the Li_2 phenotype contraindicated those locations, although leaving open the possibility that multiple nearby genes influence this trait, such as the gene for seed shattering in sorghum (*Sorghum bicolor*) involving *SpWRKY* and *YABBY* loci, which are only 300 kb apart (Lin et al., 2012; Tang et al., 2013). Further, RNAseq conducted using DP5690 (wild type) and an Li_2 mutant in two different conditions (field and greenhouse) found 13 of 14 genes on the terminal end of chromosome 18 (or D13) to be down-regulated (Naoumkina et al., 2015), providing additional independent support for our theory of deletion in this region.

Haploinsufficiency, such as we postulate to account for the genetic dominance of the mutant Li_2 allele, has been reported in plants, albeit rarely (Meinke 2013; Yuan et al., 2014). Our data corroborate prior reports of abnormal expression and incomplete penetrance of the Li_2 phenotype (An et al., 2010), resembling haploinsufficiency, variable expressivity, and incomplete penetrance associated with segmental deletion in multiple human studies (Köhn et al., 2009; Klaassen et al., 2013; Todarello et al., 2014; El Khattabi et al., 2015; Ponzi et al., 2015). The scarcity of haploinsufficiency in

Downloaded from https://academic.oup.com/plphys/article/183/1/277/6116371 by Qatar University user on 14 March 2024

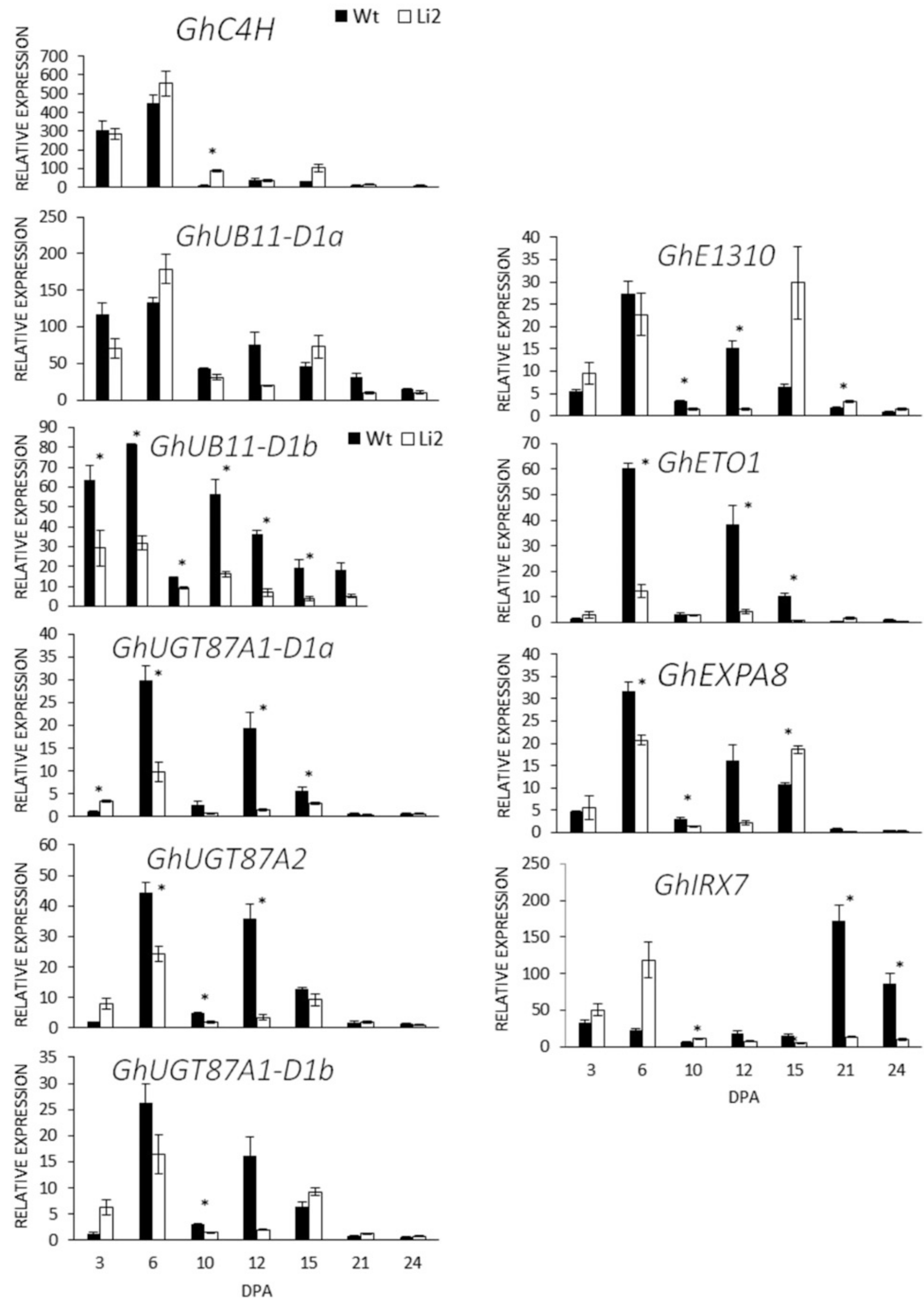

Figure 3. Gene expression analysis using RT-qPCR for candidate genes in fiber tissues from wild-type (Wt) and mutant plants. The *x* axis represents DPA. Fiber and ovule tissues were mixed for 3 and 6 DPA, and only fiber tissues were used thereafter (10, 12, 15, 21, and 24 DPA). Error bars represent the means ± SE from three biological replicates. Asterisks indicate significance at $P < 0.05$, Student's *t* test.

Downloaded from https://academic.oup.com/plphys/article/183/1/277/6116371 by Qatar University user on 14 March 2024

plants might be due to plant genomes having a greater degree of tolerance than other genomes for altered gene dosage, as plant lineages have incurred genome duplications more frequently than any other taxa known (Paterson et al., 2010). Genes with similar or related functions tend to cluster in genomes (Nei 2003; Yi et al., 2007). Deletion of part of a chromosome containing genes with related functions might have a haploinsufficient effect, as one copy of multiple genes involved in a biosynthetic pathway is removed (Thomas et al., 2006). QTL clusters for single or multiple fiber traits have been identified in cotton (Rong et al., 2007; Zhang et al., 2016) and related to closely spaced genes that are functionally diverse but coordinately regulated (Paterson et al., 2012).

Although we consider it probable that deletion of these genes causes the Li_2 phenotype, other factors may also contribute. For example, long noncoding RNAs (lncRNAs) regulate plant gene expression by manipulating transcription, splicing, and gene silencing (Chekanova 2015). Recent research has shown that lncRNAs function in fiber development (Wang et al., 2015; Hu et al., 2018; Salih et al., 2019). Deletion of lncRNAs regulating fiber-related genes could impact fiber elongation processes. Moreover, chromosome rearrangement such as deletion can impact the expression

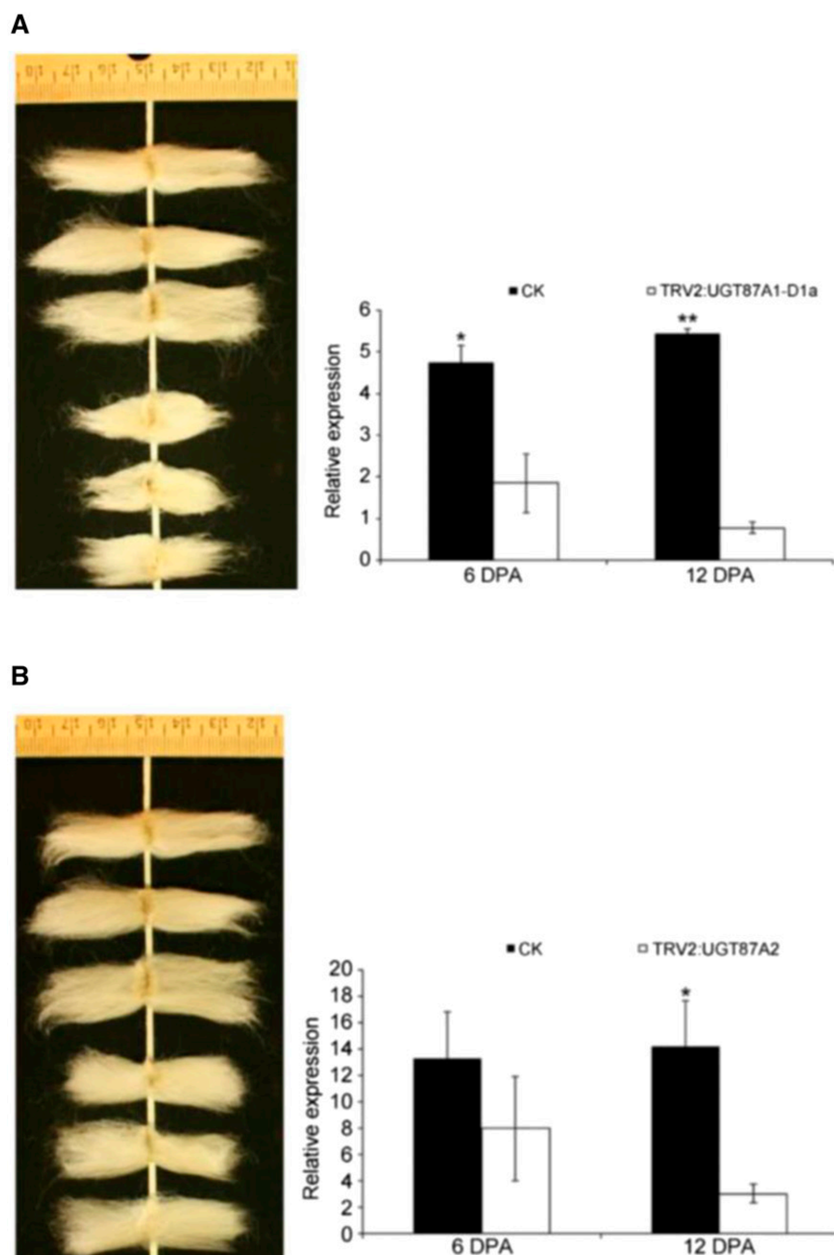

Figure 4. Reduction of fiber length by VIGS. Fiber samples from three different plants infected by TRV2:UGT87A1-D1a (A) and TRV2:UGT87A2 (B) and relative expression of the genes at different fiber development stages (6 and 12 DPA). Error bars represent the means ± SE from three biological replicates. Unpaired Student's *t* tests were performed to determine statistical significance (* $P < 0.05$ and ** $P < 0.01$).

Downloaded from https://academic.oup.com/plphys/article/183/1/277/6116371 by Qatar University user on 14 March 2024

of normal-copy neighboring genes (Merla et al., 2006; Dantas et al., 2019). This might also be a factor contributing to the Li_2 phenotype.

The failure of fibers to elongate properly in the absence of two wild-type (in this case, recessive) li_2 alleles suggests that the Dt genome of tetraploid cotton may be differentiated from its diploid progenitor by a neomorphic mutation. While diploid D-genome cotton plants do not produce lint fibers, many fiber-quality related QTLs locate in Dt genomic regions of tetraploid cottons (Rong et al., 2007; Wang et al., 2013). The fact that loss of even a single li_2 allele causes extremely short fiber suggests that homologs of these genes, including homeologs in the At genome, are functionally diverged and not able to compensate for this loss. Having a partially functional homologous gene(s) may cause abnormal expression of the wild-type phenotype and incomplete penetrance such as we observe. A careful study of gene expression using RNA sequencing in plants with such atypical phenotypes might clarify such an anomaly.

While only two genes in the Li_2 region showed a VIGS-induced fiber length phenotype, at least seven (*GhIRX7*, *GhETO1*, *GhUBE11_D1b*, *GhUGT87A1_D1a*, *GhUGT87A2*, *GhUGT87A1_D1b*, and *GhEXPA8)* of the 10 genes with annotations related to fiber development showed striking expression differences between the wild type and Li_2 mutants during fiber development stages. Due to the presence of a homeologous copy on the At part of the genome, deletion per se fails to

account for these differences, as Li_2 mutants show higher expression at some stages. These genes have diverse and potentially interrelated functions during fiber development (Fig. 5). *GhIRX7* belongs to the glycosyltransferase family 47 (Zhong and Ye 2003), including members for which disruption causes reduction in secondary wall thickness, decline in amounts of cellulose and xylan, collapse of xylem vessels, decrease in stem strength, and dwarf stature (Zhong et al., 2005; Brown et al., 2007). *GhEXPA8* belongs to the expansin superfamily of wall-loosening proteins that help with cell expansion (Sampedro and Cosgrove 2005), for which overexpression increases cotton fiber length (Bajwa et al., 2015). *GhETO1* negatively regulates 1-aminocyclopropane-1-carboxylic acid synthase (ACS; Wang et al., 2004), controlling the rate of synthesis of ethylene, which plays an important role in fiber elongation (Shi et al., 2006). *GhUBE11_D1b* encodes ubiquitin-activating enzyme E1 that catalyzes the first step of three consecutive enzymatic cascades in a ubiquitination reaction—notably, regulation of ethylene biosynthesis through ubiquitin-mediated protein degradation is assisted by *ETHYLENE-OVERPRODUCER1* (Lyzenga and Stone 2012).

GhUGT87A1_D1a, *GhUGT87A2*, and *GhUGT87A1_D1b* belong to Glycosyltransferase Family 1, the largest glycotransferase family (Yonekura-Sakakibara and Hanada 2011; Huang et al., 2015). *UGT87A1* and *UGT87A2* are nearly identical, and *UGT87A2* in Arabidopsis has been implicated in diverse traits, including abscisic acid (ABA) and ascorbic acid homeostasis, cell wall biosynthesis, and controlling the H_2O_2 level (Saint Paul, 2010; Wang et al., 2012; Li et al., 2017; Rehman et al., 2018), each active during fiber elongation (Li et al., 2007; Chaudhary et al., 2008, 2009; Hovav et al., 2008; Guo et al., 2016). Elevated concentrations of H_2O_2 occur in *G. longicalyx* (F genome) compared to *G. herbaceum* (A genome; Hovav et al., 2008) and Li_2 mutants (Hinchliffe et al., 2011; Naoumkina et al., 2013). Similarly, ABA is an important hormone for plant development, but higher concentration of ABA as observed in Li_2 mutants inhibits cotton fiber elongation and seed germination (Beasley and Ting 1973; Gokani et al., 1998; Gilbert et al., 2013;). Finally, segregation distortion in our population, with significantly reduced representation of homozygous *G. hirsutum* (Li_2 mutant) alleles, may suggest some roles of these genes in reproductive organs, seed viability, and germination.

While we tried to evaluate the candidate genes in the deletion region that seemed most likely to affect fiber development, less obvious but still salient genes may have been overlooked. For example, *OVATE FAMILY PROTEIN3* (Gorai.013G272200) belongs to a family of transcriptional repressors that regulate secondary cell wall formation through interaction with *KNOTTED ARABIDOPSIS THALIANA7* (Li et al., 2011). Similarly, a nonspecific lipid-transfer protein (Gorai.013G271900) involved in a wide range of biological processes including pathogen defense and abiotic stress response, and the transcription factor *REGULATOR OF AXILLARY MERISTEMS2* (Gorai.013G269300), which regulates shoot branching (Boutrot et al., 2008; Guo et al., 2015), may warrant future investigation.

While the cotton Li_2 mutant phenotype is essentially discrete and tentatively due to two specific genes, the finding that the mutant exhibits coordinated regulation

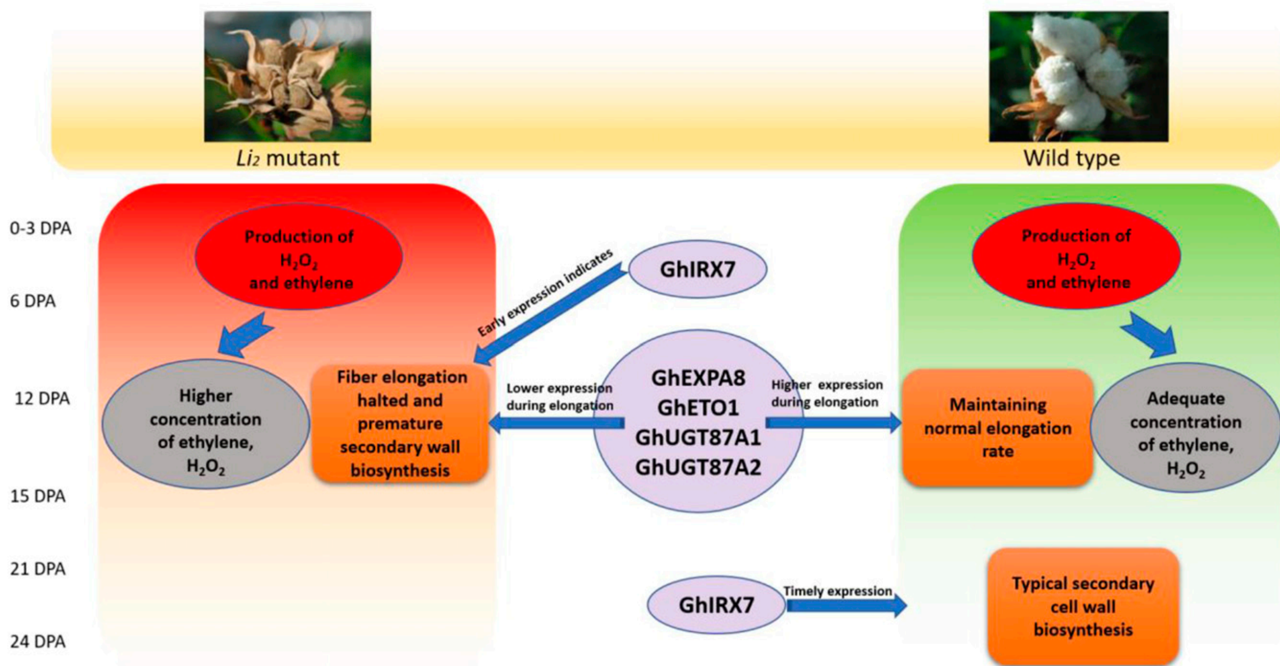

Figure 5. Inferred effect of deletion of terminal region of chromosome 18 on the fiber development process.

Downloaded from https://academic.oup.com/plphys/article/183/1/277/6116371 by Qatar University user on 14 March 2024

of nearby genes affecting processes in cell wall development and cell elongation is consistent with (albeit not proving) the hypothesis that some fiber QTLs comprise closely spaced groups of functionally diverse but coordinately regulated genes (Paterson et al., 2012). Further research using methods such as clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein 9 (Cas9) genome editing techniques to make more precise changes to single genes or their regulatory elements might provide more definitive insight into the hypothesis.

MATERIALS AND METHODS

Plant Material and Population Development

A F_2 mapping population was developed by an interspecific cross between a Li_2 mutant (TM-1; *Gossypium hirsutum*) and Pima S-7 (*Gossypium barbadense*), to obtain rich polymorphism of DNA markers. All F1 hybrids displayed the mutant phenotype, consistent with the dominant nature of Li_2. F_1 plants were selfed and a total of 1,545 F_2 individuals were grown in a greenhouse or field for this study. Noting that the Li_2 phenotype sometimes shows variation within a plant, we observed multiple bolls to determine the Li_2 trait of the F_2 plants.

Genetic Marker Development

Initial screening and identifying of linked markers near Li_2 used SSR (and expressed sequence tag (EST) markers from published genetic linkage maps (Rong et al., 2005, 2007; Wang et al., 2013) near the telomere of the long arm of chromosome 18 (Shan et al., 2016). Closely linked markers were used to select BACs from a *Gossypium raimondii* library that were sequenced, developing new markers from BAC sequences by clipping vectors, identifying SSRs, and designing primers (Freitas et al., 2008) or using conserved intron scanning polymorphism (CISP) methods (Feltus et al., 2006). A total of 144 markers were first tested in a small population to check polymorphism and proximity to Li_2. Seven markers with clear bands, and which immediately flank Li_2, were mapped in all 1,545 F_2 individuals. Genetic maps were built using JoinMap 4.1 (Van Ooijen 2011).

BAC Library Screening

For chromosome walking, cotton BAC libraries were screened with overgo probes designed from genetic markers mapping closely to the Li_2 phenotype, identifying a total of 124 BACs in five different BAC libraries that might be in the Li_2 region (Supplemental Table S1). To reduce the chances of false positive hits and to minimize sequencing-redundant BACs, we used a *G. raimondii* physical map (Lin et al., 2010), designing probes from tightly linked genetic markers and hybridizing them to a *G. raimondii* BAC library and corresponding contigs, which led to contig ctg2409 on the physical map (Lin et al., 2010). We sequenced three BACS (GR174O23, GR109E22, and GR174F21), finding markers from GR174F21 to be nearest to Li_2. To find BACs that could be sequenced to extend the terminal region of chromosome 18 and get closer to Li_2, we BLASTed 13,662 BAC end sequences (BESs) from the *G. raimondii* library (Lin et al., 2010) to identify BACs with one end sequence clearly anchored to the terminal region of chromosome D 13 (the D genome diploid chromosome corresponding to tetraploid chromosome 18), but the other end unanchored. To determine how informative these BACs were as compared to GR174F21, we performed DNA fingerprinting (Lin et al., 2010).

Validation of Deletion

A total of 12 genes were selected from across the target region, with two between CISP15 and the centromere and 10 between CISP15 and the telomere. Primers (20–24 bp with ~500 bp amplicons) were designed using Primer 3 (http://frodo.wi.mit.edu/). All amplicons had at least five SNPs differentiating between corresponding loci in the tetraploid At and Dt subgenomes. Three genotypes each from Li_2 homozygous and wild-type plants along with 12 pairs of primers were amplified using a thermal profile as follows: 95°C for 30 s (denaturation), 55°C for 30 s (annealing), and 72°C for 1 min (extension) for 36 cycles, followed by a terminal extension step at 72°C for 10 min. PCR amplicons were separated using 10% (w/v) nondenatured polyacrylamide gel. Amplicons were treated with exonuclease I and shrimp alkaline phosphatase to remove unused primers and deoxyribose nucleoside triphosphates, then amplified using BigDye mixture at 95°C for 5 min followed by 95°C for 10 s, 50°C for 5 s, and 60°C for 4 min for 30 cycles. A clean up step was performed using Sephadex plates to remove excessive primers and dyes, and amplicons were sequenced on an ABI3730 capillary sequencer. Sequence analysis was done using the Codon Code Aligner software (CodonCode Corporation) to compare the terminal end of chromosome 18 in Li_2 mutant lines to wild-type lines. We extracted At and Dt sequences from the *G. hirsutum* genome (Li et al., 2015) using a python script to compare SNPs in the sequenced regions and to validate the absence of the Dt genomic region in Li_2 mutants.

Mining for Putative Candidate Gene(s)

A 221-kb candidate region was extracted from the *G. raimondii* sequence (Paterson et al., 2012) using a python script and scanned using FGENESH to determine mRNA and protein sequences of putative genes (Salamov and Solovyev 2000) using the protein sequences of these genes for BLASTp (protein-protein blast) in the National Center for Biotechnology Information protein database to identify possible homologs and deduce possible functions. A total of 10 putative candidate genes were identified in the region with annotations suggesting functions associated with fiber elongation processes.

Comparing Gene Sequences between Different Cotton Genomes

A sequence comparison of candidate genes between elongated fiber-producing (A and AtDt (tetraploid)) and nonelongated fiber-producing (D and F) cotton genomes was performed. Reference sequence data from the D genome (*G. raimondii*; Paterson et al., 2012) and resequencing data of the A (*Gossypium herbaceum*), F (*Gossypium longicalyx*), and AtDt (*G. hirsutum* 'Acala Maxxa', *G. hirsutum* race yucatanense, and *Gossypium mustelinum*) genomes were used for this purpose, with sequences of candidate genes compared using the Burrows-Wheeler Aligner (Li and Durbin 2009) to align resequencing data to the reference *G. raimondii* (Paterson et al., 2012) and Sequence Alignment/Map (Li et al., 2009) used for SNP calling.

RNA Isolation and RT-qPCR

Fibers at different stages of cotton boll development, i.e. initiation, elongation, and secondary cell wall thickening stages, were used to examine the expression profiles of 10 genes using RT-qPCR. Tissue samples were collected at 3 (fiber + ovule mix), 6 (fiber +ovule mix), 10 (fiber), 12 (fiber), 15 (fiber), 21 (fiber) and 24 (fiber) DPA. RNA was extracted using the Purelink Plant RNA reagent kit, according to the manufacturer's instructions (Life Technologies Corporation) with a DNA digestion step using turbo DNA-free kit (Applied Biosystems). Quality of RNA was determined by electrophoresis in 1% (w/v) agarose gels and concentration was determined in a NanoDrop 2000 spectrophotometer (NanoDrop Technologies). A total of 500 ng of mRNA was used to obtain first-strand complementary DNA (cDNA) using a SuperScript III Reverse Transcriptase cDNA Synthesis kit (Thermo Fisher Scientific) according to the manufacturer's instructions. The primers designed for this experiment were gene specific and not homolog specific for At and Dt subgenomes. qPCR was carried out on a Roche LightCycler480 II instrument (Roche Diagnostics) in a 25-μL volume containing 10 ng of cDNA, 5 pM of each primer, and 25 μL of Fast SYBR Green Master Mixture (Thermo Fisher Scientific) according to the manufacturer's protocol. The PCR conditions were as follows: primary denaturation at 95°C for 20 s followed by 40 amplification cycles of 3 s at 95°C, then 30 s at 60°C. Melting curve analysis was performed to ensure there was no primer-dimer formation. Three replicate assays were performed with independently isolated RNAs. Relative expression levels of each gene are presented using the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001). The cotton *GhActin* gene was used as the positive control. Primer information of RT-qPCR for gene expression is provided in Supplemental Table S3.

VIGS

Using PCR, we amplified 351- to 599-bp fragments from cDNA for the 10 candidate genes (Supplemental Table S4) in the target region of chromosome

Downloaded from https://academic.oup.com/plphys/article/183/1/277/6116371 by Qatar University user on 14 March 2024

18. These fragments were cloned and inserted into the VIGS binary vector pYL156 (pTRV2) that was double digested with either *Eco*RI *Kpn*I, *Eco*RI *Xho*I, or *Kpn*I *Xho*I to generate inverted repeats. The vector pTRV1(pYL192), which is a helper plasmid, and cloned pTRV2 vector were introduced into *Agrobacterium tumefaciens* strain GV3101 through electroporation (Bio-Rad Gene Pulser II). For the VIGS assay, the transformed *Agrobacterium* colonies containing pTRV1 and cloned pTRV2 were inoculated into 5 mL of Luria-Bertani medium containing kanamycin (50 μg mL^{-1}) and gentamycin (25 μg mL^{-1}) and grown overnight at 28°C in an Innova 4080 shaking at a speed of 80 rpm. The next day, the cultures were transferred into a flask containing 50 mL Luria-Bertani medium with kanamycin (50 μg mL^{-1}) and gentamycin (25 μg mL^{-1}) antibiotic plus 10 mM MES and 20 μM acetosyringone and left to grow overnight at 28°C with shaking at 80 rpm (Gao et al., 2011). The culture was resuspended using infiltration buffer (10 mM $MgCl_2$, 10 mM MES, and 200 μM acetosyringone). The final OD_{600} of the culture was adjusted to 1.5. Seven-day-old seedlings with fully expanded cotyledons but no visible true leaves were infiltrated with solution made from mixtures of pTRV1 and pTRV2 vectors in a 1:1 ratio. Infected plants along with controls were grown in growth chambers under 15 h light/9 h dark cycles with light intensity of 120 μE m^{-2} S^{-1} and temperature of 24°C. We also performed BLAST searches for the sequences of genes *GhUGT87A1-D1a* and *GhUGT87A2* used for VIGS in the *Gossypium hirsutum* genome to identify any cross silencing by VIGS (Supplemental Table S5).

Statistical Analyses

Student's *t* test was performed to determine significant differences in RT-qPCR data, with $*P < 0.05$. Microsoft Excel was used to do statistical analyses and draw graphs. Excel was also used to perform χ^2 test to determine segregation distortion for the markers in the Li_2 region.

Accession Numbers

Sequence information for genes Gorai.013G270900, Gorai.013G271000, and other genes in the deletion region can be found at https://www.cottongen.org/.

Supplemental Data

The following supplemental materials are available.

Supplemental Figure S1. Identifying new BACs in Li_2 region.

Supplemental Figure S2. Validating deletion of the terminal end of chromosome 18 in Li_2 homozygous.

Supplemental Figure S3. Representative plants treated with TRV2:GRCLA1 as a visual marker for verifying efficiency of viral infection.

Supplemental Figure S4. Phylogenetic analysis and protein sequence alignment.

Supplemental Table S1. Number of BACs hit by probes derived from the Li_2 region.

Supplemental Table S2. Genes in the deletion region.

Supplemental Table S3. Names and sequences of primers used in RT-PCR expression analysis.

Supplemental Table S4. Sequences of primers used in VIGS experiment.

Supplemental Table S5. Identification of 19 or more continuous bases matches (for any cross silencing by VIGS) of GhUGT87A1-D1a (D13vigs270900) and GhUGT87A2 (D13vigs271000) sequences in the *G. hirsutum* genome.

Received December 12, 2019; accepted January 28, 2020; published February 26, 2020.

LITERATURE CITED

An C, Jenkins JN, McCarty JC, Saha S (2010) Atypical *Ligon lintless-2* phenotype in cotton. J Cotton Sci **14:** 13–16

Andres RJ, Coneva V, Frank MH, Tuttle JR, Samayoa LF, Han S-W, Kaur B, Zhu L, Fang H, Bowman DT, et al (2017) Modifications to a *LATE MERISTEM IDENTITY1* gene are responsible for the major leaf shapes of Upland cotton (*Gossypium hirsutum L.*). Proc Natl Acad Sci USA **114:** E57–E66

Bajwa KS, Shahid AA, Rao AQ, Bashir A, Aftab A, Husnain T (2015) Stable transformation and expression of *GhEXPA8* fiber expansin gene to improve fiber length and micronaire value in cotton. Front Plant Sci **6:** 838

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

Beasley C, Ting IP (1973) The effects of plant growth substances on in vitro fiber development from fertilized cotton ovules. Am J Bot **60:** 130–139

Boutrot F, Chantret N, Gautier MF (2008) Genome-wide analysis of the rice and Arabidopsis *non-specific lipid transfer protein* (*nsLtp*) gene families and identification of wheat *nsLtp* genes by EST data mining. BMC Genomics **9:** 86

Bowman DT, Van Esbroeck GA, Van't Hof J, Jividen GM (2001) Ovule fiber cell numbers in modern upland cottons. J Cotton Sci **5:** 81–83

Brown DM, Goubet F, Wong VW, Goodacre R, Stephens E, Dupree P, Turner SR (2007) Comparison of five xylan synthesis mutants reveals new insight into the mechanisms of xylan synthesis. Plant J **52:** 1154–1168

Cai C, Tong X, Liu F, Lv F, Wang H, Zhang T, Guo W (2013) Discovery and identification of a novel Ligon lintless-like mutant (*Lix*) similar to the Ligon lintless (*Li1*) in allotetraploid cotton. Theor Appl Genet **126:** 963–970

Chaudhary B, Hovav R, Flagel L, Mittler R, Wendel JF (2009) Parallel expression evolution of oxidative stress-related genes in fiber from wild and domesticated diploid and polyploid cotton (*Gossypium*). BMC Genomics **10:** 378

Chaudhary B, Hovav R, Rapp R, Verma N, Udall JA, Wendel JF (2008) Global analysis of gene expression in cotton fibers from wild and domesticated *Gossypium barbadense*. Evol Dev **10:** 567–582

Chekanova JA (2015) Long non-coding RNAs and their functions in plants. Curr Opin Plant Biol **27:** 207–216

Dantas AG, Santoro ML, Nunes N, de Mello CB, Pimenta LSE, Meloni VA, Soares DCQ, Belangero SN, Carvalheira G, Kim CA, et al (2019) Downregulation of genes outside the deleted region in individuals with 22q11.2 deletion syndrome. Hum Genet **138:** 93–103

Edgar RC (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res **32:** 1792–1797

El Khattabi L, Guimiot F, Pipiras E, Andrieux J, Baumann C, Bouquillon S, Delezoide A-L, Delobel B, Demurger F, Dessuant H, et al (2015) Incomplete penetrance and phenotypic variability of 6q16 deletions including *SIM1*. Eur J Hum Genet **23:** 1010–1018

Fang L, Tian R, Li X, Chen J, Wang S, Wang P, Zhang T (2014) Cotton fiber elongation network revealed by expression profiling of longer fiber lines introgressed with different *Gossypium barbadense* chromosome segments. BMC Genomics **15:** 838

Feltus FA, Singh HP, Lohithaswa HC, Schulze SR, Silva TD, Paterson AH (2006) A comparative genomics strategy for targeted discovery of single-nucleotide polymorphisms and conserved-noncoding sequences in orphan crops. Plant Physiol **140:** 1183–1191

Freitas PD, Martins DS, Galetti PM Jr. (2008) CID: A rapid and efficient bioinformatic tool for the detection of SSRs from genomic libraries. Mol Ecol Resour **8:** 107–108

Gao X, Britt RC Jr., Shan L, He P (2011) Agrobacterium-mediated virus-induced gene silencing assay in cotton. J Vis Exp **54:** e2938

Gilbert MK, Bland JM, Shockey JM, Cao H, Hinchliffe DJ, Fang DD, Naoumkina M (2013) A transcript profiling approach reveals an abscisic acid-specific glycosyltransferase (*UGT73C14*) induced in developing fiber of Ligon lintless-2 mutant of cotton (*Gossypium hirsutum L.*). PLoS One **8:** e75268

Gokani SJ, Kumar R, Thaker VS (1998) Potential role of abscisic acid in cotton fiber and ovule development. J Plant Growth Regul **17:** 1–5

Griffee F, Ligon L (1929) Occurrence of "lintless" cotton plants and inheritance of character "lintless". J Am Soc Agron **21:** 711–717

Guo D, Zhang J, Wang X, Han X, Wei B, Wang J, Li B, Yu H, Huang Q, Gu H, et al (2015) The *WRKY* transcription factor *WRKY71/EXB1* controls shoot branching by transcriptionally regulating *RAX* genes in *Arabidopsis*. Plant Cell **27:** 3112–3127

Guo K, Du X, Tu L, Tang W, Wang P, Wang M, Liu Z, Zhang X (2016) Fibre elongation requires normal redox homeostasis modulated by

Downloaded from https://academic.oup.com/plphys/article/183/1/277/6116371 by Qatar University user on 14 March 2024

cytosolic ascorbate peroxidase in cotton (*Gossypium hirsutum*). J Exp Bot **67:** 3289–3301

Hinchliffe DJ, Turley RB, Naoumkina M, Kim HJ, Tang Y, Yeater KM, Li P, Fang DD (2011) A combined functional and structural genomics approach identified an EST-SSR marker with complete linkage to the Ligon lintless-2 genetic locus in cotton (*Gossypium hirsutum L.*). BMC Genomics **12:** 445

Hovav R, Udall JA, Chaudhary B, Hovav E, Flagel L, Hu G, Wendel JF (2008) The evolution of spinnable cotton fiber entailed prolonged development and a novel metabolism. PLoS Genet **4:** e25

Hu H, Wang M, Ding Y, Zhu S, Zhao G, Tu L, Zhang X (2018) Transcriptomic repertoires depict the initiation of lint and fuzz fibres in cotton (*Gossypium hirsutum L.*). Plant Biotechnol J **16:** 1002–1012

Huang J, Pang C, Fan S, Song M, Yu J, Wei H, Ma Q, Li L, Zhang C, Yu S (2015) Genome-wide analysis of the family 1 glycosyltransferases in cotton. Mol Genet Genomics **290:** 1805–1818

Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci **8:** 275–282

Kim HJ, Triplett BA (2001) Cotton fiber growth in planta and in vitro. Models for plant cell elongation and cell wall biogenesis. Plant Physiol **127:** 1361–1366

Klaassen P, Duijff S, Swanenburg de Veye H, Vorstman J, Beemer F, Sinnema G (2013) Behavior in preschool children with the 22q11.2 deletion syndrome. Am J Med Genet A **161A:** 94–101

Köhn L, Bowne SJ, S Sullivan L, Daiger SP, Burstedt MS, Kadzhaev K, Sandgren O, Golovleva I (2009) Breakpoint characterization of a novel ~59 kb genomic deletion on 19q13.42 in autosomal-dominant retinitis pigmentosa with incomplete penetrance. Eur J Hum Genet **17:** 651–655

Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol **33:** 1870–1874

Lee JJ, Woodward AW, Chen ZJ (2007) Gene expression changes and early events in cotton fibre development. Ann Bot **100:** 1391–1401

Li D-D, Ruan X-M, Zhang J, Wu Y-J, Wang X-L, Li X-B (2013) Cotton plasma membrane intrinsic protein 2s (*PIP2s*) selectively interact to regulate their water channel activities and are required for fibre development. New Phytol **199:** 695–707

Li E, Wang S, Liu Y, Chen JG, Douglas CJ (2011) OVATE FAMILY PROTEIN4 (OFP4) interaction with *KNAT7* regulates secondary cell wall formation in *Arabidopsis thaliana*. Plant J **67:** 328–341

Li F, Fan G, Lu C, Xiao G, Zou C, Kohel RJ, Ma Z, Shang H, Ma X, Wu J, et al (2015) Genome sequence of cultivated Upland cotton (*Gossypium hirsutum* TM-1) provides insights into genome evolution. Nat Biotechnol **33:** 524–530

Li F, Fan G, Wang K, Sun F, Yuan Y, Song G, Li Q, Ma Z, Lu C, Zou C, et al (2014) Genome sequence of the cultivated cotton *Gossypium arboreum*. Nat Genet **46:** 567–572

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics **25:** 1754–1760

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R; 1000 Genome Project Data Processing Subgroup (2009) The sequence alignment/map format and SAMtools. Bioinformatics **25:** 2078–2079

Li HB, Qin YM, Pang Y, Song WQ, Mei WQ, Zhu YX (2007) A cotton ascorbate peroxidase is involved in hydrogen peroxide homeostasis during fibre cell development. New Phytol **175:** 462–471

Li P, Li YJ, Wang B, Yu HM, Li Q, Hou BK (2017) The Arabidopsis *UGT87A2*, a stress-inducible family 1 glycosyltransferase, is involved in the plant adaptation to abiotic stresses. Physiol Plant **159:** 416–432

Lin L, Pierce GJ, Bowers JE, Estill JC, Compton RO, Rainville LK, Kim C, Lemke C, Rong J, Tang H, et al (2010) A draft physical map of a D-genome cotton species (*Gossypium raimondii*). BMC Genomics **11:** 395

Lin Z, Li X, Shannon LM, Yeh C-T, Wang ML, Bai G, Peng Z, Li J, Trick HN, Clemente TE, et al (2012) Parallel domestication of the *Shattering1* genes in cereals. Nat Genet **44:** 720–724

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_T}$ method. methods **25:** 402–408

Lyzenga WJ, Stone SL (2012) Regulation of ethylene biosynthesis through protein degradation. Plant Signal Behav **7:** 1438–1442

Meinke DW (2013) A survey of dominant mutations in *Arabidopsis thaliana*. Trends Plant Sci **18:** 84–91

Merla G, Howald C, Henrichsen CN, Lyle R, Wyss C, Zabot MT, Antonarakis SE, Reymond A (2006) Submicroscopic deletion in patients with Williams-Beuren syndrome influences expression levels of the nonhemizygous flanking genes. Am J Hum Genet **79:** 332–341

Naoumkina M, Hinchliffe DJ, Turley RB, Bland JM, Fang DD (2013) Integrated metabolomics and genomics analysis provides new insights into the fiber elongation process in Ligon lintless-2 mutant cotton (*Gossypium hirsutum L.*). BMC Genomics **14:** 155

Naoumkina M, Thyssen GN, Fang DD (2015) RNA-seq analysis of short fiber mutants Ligon-lintless-1 (*Li 1*) and -2 (Li_2) revealed important role of aquaporins in cotton (*Gossypium hirsutum L.*) fiber elongation. BMC Plant Biol **15:** 65

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

Nei M (2003) Genome evolution: Let's stick together. Heredity **90:** 411–412

Osmani SA, Bak S, Møller BL (2009) Substrate specificity of plant UDP-dependent glycosyltransferases predicted from crystal structures and homology modeling. Phytochemistry **70:** 325–347

Pang CY, Wang H, Pang Y, Xu C, Jiao Y, Qin YM, Western TL, Yu SX, Zhu YX (2010) Comparative proteomics indicates that biosynthesis of pectic precursors is important for cotton fiber and *Arabidopsis* root hair elongation. Mol Cell Proteomics **9:** 2019–2033

Paterson AH, Freeling M, Tang H, Wang X (2010) Insights from the comparison of plant genome sequences. Annu Rev Plant Biol **61:** 349–372

Paterson AH, Wendel JF, Gundlach H, Guo H, Jenkins J, Jin D, Llewellyn D, Showmaker KC, Shu S, Udall J, et al (2012) Repeated polyploidization of *Gossypium* genomes and the evolution of spinnable cotton fibres. Nature **492:** 423–427

Ponzi E, Asaro A, Orteschi D, Genuardi M, Zollino M, Gurrieri F (2015) Variable expressivity of a familial 1.9 Mb microdeletion in 3q28 leading to haploinsufficiency of TP63: Refinement of the critical region for a new microdeletion phenotype. Eur J Med Genet **58:** 400–405

Rehman HM, Nawaz MA, Shah ZH, Ludwig-Müller J, Chung G, Ahmad MQ, Yang SH, Lee SI (2018) Comparative genomic and transcriptomic analyses of Family-1 UDP glycosyltransferase in three *Brassica* species and *Arabidopsis* indicates stress-responsive regulation. Sci Rep **8:** 1875

Rong J, Feltus FA, Waghmare VN, Pierce GJ, Chee PW, Draye X, Saranga Y, Wright RJ, Wilkins TA, May OL, et al (2007) Meta-analysis of polyploid cotton QTL shows unequal contributions of subgenomes to a complex network of genes and gene clusters implicated in lint fiber development. Genetics **176:** 2577–2588

Rong J, Pierce GJ, Waghmare VN, Rogers CJ, Desai A, Chee PW, May OL, Gannaway JR, Wendel JF, Wilkins TA, et al (2005) Genetic mapping and comparative analysis of seven mutants related to seed fiber development in cotton. Theor Appl Genet **111:** 1137–1146

Salih H, Gong W, He S, Xia W, Odongo MR, Du X (2019) Long non-coding RNAs and their potential functions in Ligon-lintless-1 mutant cotton during fiber development. BMC Genomics **20:** 661

Saint Paul Vv (2010) Stress inducible glycosyltransferases in Arabidopsis thaliana and their impact on plant metabolism and defense mechanisms. PhD thesis. Ludwig-Maximilians Universität München, Munich, Germany

Salamov AA, Solovyev VV (2000) Ab initio gene finding in *Drosophila* genomic DNA. Genome Res **10:** 516–522

Sampedro J, Cosgrove DJ (2005) The expansin superfamily. Genome Biol **6:** 242

Shan C-M, Shangguan X-X, Zhao B, Zhang X-F, Chao LM, Yang C-Q, Wang L-J, Zhu H-Y, Zeng Y-D, Guo W-Z, et al (2014) Control of cotton fibre elongation by a homeodomain transcription factor *GhHOX3*. Nat Commun **5:** 5519

Shan W, Jiang Y, Han J, Wang K (2016) Comprehensive cytological characterization of the *Gossypium hirsutum* genome based on the development of a set of chromosome cytological markers. Crop J **4:** 256–265

Shi YH, Zhu SW, Mao XZ, Feng JX, Qin YM, Zhang L, Cheng J, Wei LP, Wang ZY, Zhu YX (2006) Transcriptome profiling, molecular biological, and physiological studies reveal a major role for ethylene in cotton fiber cell elongation. Plant Cell **18:** 651–664

Tang H, Cuevas HE, Das S, Sezen UU, Zhou C, Guo H, Goff VH, Ge Z, Clemente TE, Paterson AH (2013) Seed shattering in a wild sorghum is conferred by a locus unrelated to domestication. Proc Natl Acad Sci USA **110:** 15824–15829

Tang W, Tu L, Yang X, Tan J, Deng F, Hao J, Guo K, Lindsey K, Zhang X (2014) The calcium sensor *GhCaM7* promotes cotton fiber elongation by

Downloaded from https://academic.oup.com/plphys/article/183/1/277/6116371 by Qatar University user on 14 March 2024

modulating reactive oxygen species (ROS) production. New Phytol **202:** 509–520

Thomas BC, Pedersen B, Freeling M (2006) Following tetraploidy in an *Arabidopsis* ancestor, genes were removed preferentially from one homeolog leaving clusters enriched in dose-sensitive genes. Genome Res **16:** 934–946

Thyssen GN, Fang DD, Turley RB, Florane C, Li P, Naoumkina M (2014) Next generation genetic mapping of the Ligon-lintless-2 (Li_2) locus in upland cotton (*Gossypium hirsutum* L.). Theor Appl Genet **127:** 2183–2192

Todarello G, Feng N, Kolachana BS, Li C, Vakkalanka R, Bertolino A, Weinberger DR, Straub RE (2014) Incomplete penetrance of *NRXN1* deletions in families with schizophrenia. Schizophr Res **155:** 1–7

Wan Q, Guan X, Yang N, Wu H, Pan M, Liu B, Fang L, Yang S, Hu Y, Ye W, et al (2016) Small interfering RNAs from bidirectional transcripts of *GhMML3_A12* regulate cotton fiber development. New Phytol **210:** 1298–1310

Wang B, Jin SH, Hu HQ, Sun YG, Wang YW, Han P, Hou BK (2012) *UGT87A2*, an *Arabidopsis* glycosyltransferase, regulates flowering time via FLOWERING LOCUS C. New Phytol **194:** 666–675

Wang KL-C, Yoshida H, Lurin C, Ecker JR (2004) Regulation of ethylene gas biosynthesis by the *Arabidopsis ETO1* protein. Nature **428:** 945–950

Wang M, Yuan D, Tu L, Gao W, He Y, Hu H, Wang P, Liu N, Lindsey K, Zhang X (2015) Long noncoding RNAs and their proposed functions in fibre development of cotton (*Gossypium spp.*). New Phytol **207:** 1181–1197

Wang Z, Zhang D, Wang X, Tan X, Guo H, Paterson AH (2013) A whole-genome DNA marker map for cotton based on the D-genome sequence of Gossypium raimondii L. G3 **3:** 1759–1767

Yang Z, Zhang C, Yang X, Liu K, Wu Z, Zhang X, Zheng W, Xun Q, Liu C, Lu L, et al (2014) *PAG1*, a cotton brassinosteroid catabolism gene, modulates fiber elongation. New Phytol **203:** 437–448

Yi G, Sze S-H, Thon MR (2007) Identifying clusters of functionally related genes in genomes. Bioinformatics **23:** 1053–1060

Yonekura-Sakakibara K, Hanada K (2011) An evolutionary view of functional diversity in family 1 glycosyltransferases. Plant J **66:** 182–193

Yuan L, Dou Y, Kianian SF, Zhang C, Holding DR (2014) Deletion mutagenesis identifies a haploinsufficient role for γ-zein in opaque2 endosperm modification. Plant Physiol **164:** 119–130

Zhang S-W, Zhu X-F, Feng L-C, Gao X, Yang B, Zhang T-Z, Zhou B-L (2016) Mapping of fiber quality QTLs reveals useful variation and footprints of cotton domestication using introgression lines. Sci Rep **6:** 31954

Zhang T, Hu Y, Jiang W, Fang L, Guan X, Chen J, Zhang J, Saski CA, Scheffler BE, Stelly DM, et al (2015) Sequencing of allotetraploid cotton (*Gossypium hirsutum L.* acc. TM-1) provides a resource for fiber improvement. Nat Biotechnol **33:** 531–537

Zhong R, Peña MJ, Zhou G-K, Nairn CJ, Wood-Jones A, Richardson EA, Morrison WH III, Darvill AG, York WS, Ye Z-H (2005) *Arabidopsis fragile fiber8*, which encodes a putative glucuronyltransferase, is essential for normal secondary wall synthesis. Plant Cell **17:** 3390–3408

Zhong R, Ye Z-H (2003) Unraveling the functions of glycosyltransferase family 47 in plants. Trends Plant Sci **8:** 565–568

Downloaded from https://academic.oup.com/plphys/article/183/1/277/6116371 by Qatar University user on 14 March 2024