The $Li_2$ Mutation Results in Reduced Subgenome Expression Bias in Elongating Fibers of Allotetraploid Cotton ($Gossypium hirsutum$ L.)

Marina Naoumkina1*, Gregory Thyssen1, David D. Fang1, Doug J. Hinchliffe2, Christopher Florane1, Kathleen M. Yeater3, Justin T. Page4, Joshua A. Udall4

1 Cotton Fiber Bioscience Research Unit, USDA-ARS, Southern Regional Research Center, New Orleans, Louisiana, United States of America, 2 Cotton Chemistry & Utilization Research Unit, USDA-ARS, Southern Regional Research Center, New Orleans, Louisiana, United States of America, 3 USDA-ARS-Southern Plains Area, College Station, Texas, United States of America, 4 Plant and Wildlife Science Department, Brigham Young University, Provo, Utah, United States of America

Abstract

Next generation sequencing (RNA-seq) technology was used to evaluate the effects of the Ligon lintless-2 ($Li_2$) short fiber mutation on transcriptomes of both subgenomes of allotetraploid cotton ($Gossypium hirsutum$ L.) as compared to its near-isogenic wild type. Sequencing was performed on 4 libraries from developing fibers of $Li_2$ mutant and wild type near-isogenic lines at the peak of elongation followed by mapping and PolyCat categorization of RNA-seq data to the reference D$_5$ genome ($G. raimondii$) for homeologous gene expression analysis. The majority of homeologous genes, 83.6% according to the reference genome, were expressed during fiber elongation. Our results revealed: 1) approximately two times more genes were induced in the A$_1$ subgenome comparing to the D$_7$ subgenome in wild type and mutant fiber; 2) the subgenome expression bias was significantly reduced in the $Li_2$ fiber transcriptome; 3) $Li_2$ had a significantly greater effect on the D$_7$ than on the A$_1$ subgenome. Transcriptional regulators and cell wall homeologous genes significantly affected by the $Li_2$ mutation were reviewed in detail. This is the first report to explore the effects of a single mutation on homeologous gene expression in allotetraploid cotton. These results provide deeper insights into the evolution of allotetraploid cotton gene expression and cotton fiber development.

Citation: Naoumkina M, Thyssen G, Fang DD, Hinchliffe DJ, Florane C, et al. (2014) The $Li_2$ Mutation Results in Reduced Subgenome Expression Bias in Elongating Fibers of Allotetraploid Cotton ($Gossypium hirsutum$ L.). PLoS ONE 9(3): e90830. doi:10.1371/journal.pone.0090830

Editor: Tianzhen Zhang, Nanjing Agricultural University, China

Received September 30, 2013; Accepted February 4, 2014; Published March 5, 2014

This is an open-access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the Creative Commons CC0 public domain dedication.

Funding: This research was funded by United States Department of Agriculture-Agricultural Research Service project number 6435-21000-017-00D and Cotton Incorporated project number 58-6435-2-663. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: Co-author David D. Fang is a PLOS ONE Editorial Board member. The authors received funding from a commercial source, Cotton Incorporated. This does not alter the authors’ adherence to all the PLOS ONE policies on sharing data and materials.

* E-mail: marina.naoumkina@ars.usda.gov

Introduction

Cotton is the major source of natural fibers used in the textile industry. There are four cultivated species: AA genome diploids, $Gossypium arboresum$ L. and $G. herbaceum$ L.; and AADD genome allotetraploids, $G. hirsutum$ L. and $G. barbadense$ L. Upland cotton ($G. hirsutum$) represents about 95% of world cotton production [1]. Allotetraploid species originated around 1–2 million years ago from inter-specific hybridization between an AA-genome diploid native to Africa and Mexican DD-genome diploid [1,2].

Cotton fibers are single-celled trichomes that emerge from the ovule epidermal cells. About 25–30% of the seed epidermal cells differentiate into spinnable fibers [3]. Fiber length ranges from short (fuzz <6 mm) to long (lint). Lint fibers of economically important $G. hirsutum$ generally grow up to about 30–40 mm in length. Cotton fiber development undergoes four distinctive but overlapping stages: initiation, elongation, secondary cell wall biosynthesis, and maturation [4]. The rate and duration of each developmental stage is important to the quality attributes of the mature fiber. Cell elongation is crucial for fiber length, whereas secondary cell wall thickening is important for fiber fineness and strength.

Cotton fiber mutants are useful tools to elucidate biological processes of cotton fiber development. A cotton plant with abnormally short lint fibers was discovered in a breeding nursery of the Texas Agricultural Experiment Station in 1984. This mutant had short lint fibers (<6 mm) visually similar to those produced by Ligon lintless-1 ($Li_1$); however, unlike the stunted and deformed vegetative morphology caused by the $Li_1$ mutation, this fiber mutant had normal vegetative growth. The trait was controlled by one dominant gene named Ligon lintless-2 ($Li_2$) [5]. This gene was mapped to chromosome 18 (D$_7$ subgenome of $G. hirsutum$) using several approaches [6–8]. In a fiber developmental study, Kohel et al. [9] observed that elongation is restricted in $Li_2$ fibers, however secondary wall development proceeds normally in proportion to fiber length. Two near-isogenic lines (NILs) of $Li_2$ with the Upland cotton variety DP5690 were developed in a backcross program at Stoneville, MS [6]. Morphological evaluation of developing fibers did not reveal apparent differences between WT and $Li_2$ NILs during initiation or early elongation up to 5 days post-anthesis (DPA). Transcript and metabolite evaluations revealed significant changes in biological processes associated with cell expansion in the $Li_2$
mutant line at peak of fiber elongation, including reactive oxygen species, hormone homeostasis, nitrogen metabolism, carbohydrate biosynthesis, cell wall biogenesis, and cytoskeleton [6,10]. Therefore, the Li2 mutation can be considered as a factor affecting cotton fiber elongation process, making it an excellent model system to study cotton fiber elongation.

In previous reports, we used microarray techniques to investigate global gene expression in Li2 NILs [6,10]. However, by using the genome sequence of G. raimondii [11], RNA-seq can provide a more comprehensive and accurate transcriptome analysis based on the reference DNA sequences [12]. RNA-seq offers a larger dynamic range of quantification, reduced technical variability, and higher accuracy for distinguishing and quantifying expression levels of homologous copies than DNA microarrays [12]. Because of the limited sequence divergence between the AT and DT subgenomes in cotton, a pipeline was developed to map and categorize RNA-seq reads as originating from the AT or DT subgenomes [14].

In the present study we compared quantitative gene expression levels of RNA-seq data between developing fibers of Li2 and its WT NILs. We investigated the Li2 mutation’s effect on global transcriptional changes in subgenomes and on the functional distribution of homologous genes during fiber elongation. These results provide deeper insights into the evolution of allotetraploid cotton gene expression.

**Results**

**RNA-seq of Wild Type and Li2 Developing Fibers at Peak of Elongation**

Considering the cost of deep sequencing, only one time point, at the peak of elongation, was selected for RNA-seq analysis, including two biological replicates for wild type and mutant NILs. The time points 8 and 12 days post-anthesis (DPA) represent peak rates of fiber elongation. The time point 8 DPA was selected because: 1) our earlier research revealed significant transcript and metabolite changes between the Li2 and wild type NILs during this time of fiber development [6,10]; 2) the transcript level of the elongation stage-related gene GhEsp1 significantly decreased in Li2 mutant fiber at 8 DPA [6].

A total of 639 million reads (each 101 bp in length) from 4 libraries were obtained by paired-end Illumina sequencing. Approximately 2.3% more reads were obtained from Li2 than wild type fiber transcriptomes. From 84.4% to 90.2% of the reads was unable to be categorized as originating from either the AT or DT subgenomes or as a chimeric read (A-reads, D-reads, and X-reads, respectively; [14]). If a read did not overlap a homeologous SNP position, the read was unable to be categorized as originating from either the AT or DT subgenome (Table 1). Notably, more reads from each library were aligned to the AT subgenome than to the DT subgenome. Among the 37,223 genes on the 13 chromosomes of the G. raimondii genome, 34,892 genes (93%) had at least one mapped read from developing fibers at peak of elongation (Table 2).

**Differential Gene Expression in Developing Fibers**

Counts of mapped reads were evaluated in wild type and mutant fiber transcriptomes. Genes were considered to be expressed if they had ≥10 reads mapped in one sample. Genes that were not considered to be expressed were not included in further analyses. Approximately 3% more expressed genes were detected in Li2 than in wild type. Of the 37,223 genes on the 13 chromosomes of the G. raimondii DT reference genome 29,603 (79.5%) genes were expressed in wild type and 30,842 (82.9%) genes were expressed in Li2 fiber (Table 2).

Many genes had altered expression levels as a result of the Li2 mutation (Table S1 in File S1). Some genes were expressed in one treatment (such as Li2) but not the other treatment. For example, expressions of genes annotated as SAUR-like auxin-responsive protein (Gorai.005G257000), bHLH (Gorai.003G034700) and NAC domain transcription factor (Gorai.009G170700) were only detected in wild-type fiber. Cytokinin response factor 6 (Gorai.007G105600), UGT73C14 (Gorai.002G107900), cystein proteinase (Gorai.007G329600), MYB-like 102 (Gorai.012G132200) and WRKY transcription factor (Gorai.009G157900) were only detected in Li2 fiber. The majority of these genes have not yet been functionally characterized in cotton, except of glycosyltransferase UGT73C14, which has been shown recently to be involved in ABA homeostasis [15].

The quantitative levels of mapped reads were evaluated for differential expression between elongating fibers of Li2 and wild type NILs. A gene-by-gene ANOVA determined that 7,163 of 31,114 expressed genes were differentially expressed (FDR corrected p-value <0.05) and had ≥2-fold difference in at least one of the following comparisons: fiber type (Li2 versus wild type), AT/DT subgenomes, and combinations of these factors (statistical data for significantly regulated genes are provided in Data S1). The highest numbers of significantly differentially expressed genes were identified between the AT and DT subgenomes in wild type and Li2 whereas approximately 3 times fewer differentially expressed genes were detected between fiber type comparisons (Figure 1A). Of the 30,842 expressed genes, 3,967 (LiA/LiD, 12.9%) were differentially expressed between fiber type comparisons (Figure 1B). Of the 29,603 expressed homologous pairs in wild type, 4,578 (wtA/wtD, 15.5%) showed significantly different expression level between subgenomes; whereas in mutant fiber of the 30,842 expressed genes, 3,967 (LiA/LiD, 12.9%) were differentially expressed between subgenomes. Therefore, the homolog expression bias was significantly (p-value <0.0001; Chi square) reduced in Li2 fiber transcriptome.

In general, the AT subgenome contributed more differentially expressed genes to fiber transcriptome than did the DT subgenome. Approximately two times more genes were differentially expressed in the AT subgenome compared to the DT subgenome in wild type (AT - 2574 vs. DT - 1393) and mutant fiber (AT - 2574 vs. DT - 1393). Comparison between fiber types showed more genes were upregulated in Li2 versus wild type in both subgenomes (Figure 1B). It should be noted that only about 38% (583 genes out of 1,536 in AT and 1,511 in DT subgenome) of significantly regulated genes between mutant and wild type overlapped between subgenomes (Figure 1A).

**Mutation Effects on Transcriptome of AT and DT Subgenomes of Allotetraploid G. hirsutum**

The effect of Li2 mutation on the transcriptome of each subgenome was evaluated. The genes significantly (FDR corrected p-value <0.05) up-regulated (≥2-fold) in one subgenome of wild type were considered to have biased expression. Of the 2,958 AT biased genes, 26.5% (784) had significantly changed the expression levels in both subgenomes of Li2 as a result of mutation. However, of the 1,620 DT biased genes, 35.9% (582) had significantly
changed the expression levels in both subgenomes of the Li2 mutant (Figure 2). Therefore Li2 has a significantly greater effect ($p$-value <0.0001; Chi square) on DT biased genes than AT biased genes. Importantly, the majority of biased genes had significantly reduced expression levels (3.6% and 12.4%), whereas only a small portion of genes increased expression levels (1.6% and 2.7%) in the mutant. However, more genes, which were down-regulated in homoeologous subgenome in wild type, had increased expression levels in the mutant fiber: 11.9% and 14% were up-regulated, whereas 4.4% and 6.8% were down-regulated (Figure 2).

Furthermore, a few homeolog pairs had reciprocal expression biases between two subgenomes as a result of mutation. Expression levels for three of these genes were tested by RT-qPCR across eight developmental time points from DOA to 20 DPA, representing initiation, elongation, and beginning of secondary cell wall biosynthesis stages (Figure 3). Interestingly, the direction of expression bias changed between developmental stages in these three genes. For example, expression of homeolog pair was biased in favor of the DT subgenome for G. raimondii at initiation (1 DPA), but switched to favor the AT subgenome at elongation (5–16 DPA) in wild type developing fibers, whereas expression was biased in favor of the DT subgenome across all evaluated time points in mutant fibers. These results demonstrate that the Li2 mutation had a greater effect on the DT subgenome and also influenced direction of expression bias for some genes across developmental stages.

**Table 1. Results of mapping reads.**

| Library         | Li2 BR1 | Li2 BR2 | WT BR1 | WT BR2 |
|-----------------|---------|---------|--------|--------|
|                  | count   | %       | count  | %       | count   | %       | count  | %       |
| Number reads     | 155,057,542 | 100.0  | 168,114,870 | 100.0  | 149,831,738 | 100.0  | 165,935,842 | 100.0  |
| A reads          | 45,382,923  | 29.3   | 51,937,609   | 30.9   | 47,115,502   | 31.4   | 52,072,768   | 31.4   |
| D reads          | 36,234,551  | 23.4   | 41,617,562   | 24.8   | 37,675,249   | 25.1   | 41,512,224   | 25.0   |
| X reads          | 7,972,844  | 5.1    | 9,130,944    | 5.4    | 8,068,891    | 5.4    | 8,896,820    | 5.4    |
| N reads          | 41,252,612 | 26.6   | 46,157,152   | 27.5   | 42,291,972   | 28.2   | 46,818,360   | 28.2   |
| Mapped total     | 130,842,930| 84.4   | 148,843,267 | 88.5   | 135,151,614 | 90.2   | 149,300,172 | 90.0   |

doi:10.1371/journal.pone.0090830.t001

**Table 2. Count of expressed genes.**

| Sample                  | Li2 BR 1 | Li2 BR 2 | Li2 Total | WT BR1 | WT BR2 | WT Total | Total |
|-------------------------|----------|----------|-----------|--------|--------|----------|-------|
| Expressed (≥1)          | 32,413   | 31,845   | 33,533    | 32,170 | 32,481 | 33,460   | 34,692 |
| % of annotated genes of G. raimondii | 87.1 | 85.6 | 90.1 | 86.4 | 87.3 | 89.9 | 93.2 |
| Expressed (≥10)         | 29,919   | 29,820   | 30,842    | 28,796 | 29,041 | 29,603   | 31,114 |
| % of annotated genes of G. raimondii | 80.4 | 80.1 | 82.9 | 77.4 | 78.0 | 79.5 | 83.6 |

doi:10.1371/journal.pone.0090830.t002

The distribution of genes from the AT and DT subgenomes with significantly changed expression levels in the mutant were categorized into MapMan functional categories (Figure 4; Table S2 in File S1). Relative gene frequencies in functional categories were represented in percents of biased genes in each subgenome (2,958 AT biased genes and 1,620 DT biased genes). Most functional categories were biased in favor of the DT subgenome with the exception of photosynthesis and redox, which only contained AT homeologs. Two functional categories were significantly ($p$-value <0.05; Fisher’s exact test) biased in enrichment among DT biased genes: secondary metabolism and stress (Table S2 in File S1). These results demonstrate that different biological processes were unequally affected by Li2 mutation.

**Transcriptional regulators.** Transcriptional regulators (TRs) were identified in the G. raimondii genome based on similarity to Arabidopsis TRs and categorized into 76 families. Among them, 229 homeolog pair were AT biased and 111 were DT biased in elongating cotton fibers. Of the 229 AT biased TRs, 21 (9.2%) of them changed transcription level, whereas of the 111 DT biased TRs 14 (12.6%) of them changed transcription level (Table 3), but this difference was statistically insignificant. Expression levels for the majority of subgenome biased homeologs decreased as the result of Li2 mutation. However, six TRs (including both homeologs) had increased expression levels in mutant fibers. Three classes of TRs were the most abundant, including Aux/IAA (6 members), bHLH (5 members) and MYB (3 members). Interestingly, two of the three members of MYB TRs had increased expression due to mutation.

**Cell wall.** In the cell wall functional category, 60 homeologs were AT biased and 40 were DT biased in fiber transcriptome. Ten (16.7%) of the AT biased homeologs changed expression levels; whereas 12 (30%) of the DT biased homeologs changed expression levels (Table 4), indication a higher, but statistically insignificant effect on DT biased homeologs. Interestingly, more DT homeologs (11) than AT homeologs (4) increased transcript levels as a result of the Li2 mutation. Genes encoding enzymes involved in polysac-
charide degradation (14 genes) and cell wall proteins (9 genes) were the most abundant classes.

Validation of Illumina RNA-seq Expression and Subgenome Specific Categorization of Reads by RT-qPCR Analysis

To test the reliability of Illumina sequencing and SNP-based categorization of reads to the \( A_T \) or \( D_T \) subgenome of allotetraploid \( G. \) hirsutum, RT-qPCR analysis was performed for a subset of 8 genes (selected from Table S1 in File S1) expressed only in WT or \( Li_2 \) NILs, and for a subset of 11 genes (selected from Tables 3 and 4) that showed subgenome biased expression. Overall, the results of RT-qPCR analysis were consistent with results of RNA-seq analysis for 19 selected genes (Figures S1 and S2). RT-qPCR analysis confirmed silencing or activation of the \( Li_2 \) mutation for the subset of 8 genes (Figure S1). Correlation analysis of the expression patterns revealed strong correlations between RNA-seq and RT-qPCR data. In the subset of 11 subgenome biased genes, 7 genes showed 100% correlation (\( p \)-value <0.05) and 4 genes showed 99% correlation (\( p \)-value > 0.05; Figure S2).

Discussion

Our results demonstrate that the \( A_T \) subgenome in general contributed approximately two times more significantly induced genes to the fiber transcriptome than the \( D_T \) subgenome; however, the \( Li_2 \) mutation had greater effects on the \( D_T \) subgenome than the \( A_T \) subgenome.

Global Transcript Changes in Subgenomes of \( G. \) hirsutum Following \( Li_2 \) Mutation

The role of the \( A_T \) and \( D_T \) subgenomes in determination of fiber quality in allotetraploid cotton has been extensively discussed in the literature. Allopolyploidization resulted in significant improvements in the desirable agronomic fiber traits in the allotetraploid species in comparison with the diploid progenitors [17,18]. The first evidence showing that QTLs for fiber quality (including length, strength and fineness) were associated with DNA markers mapped to the \( D_T \) subgenome rather than the \( A_T \) subgenome was published by Paterson’s group [18]. Review of numerous QTLs published from 1998 to 2007 confirmed the observation that the \( D_T \) subgenome plays a larger role in genetic control of fiber traits [19]. A microarray study published by Wendel’s group found that the homeolog expression in \( G. \) hirsutum was biased in favor of the \( D_T \) subgenome in fiber cells [16]. Similar results were reported by Lacape and coauthors utilizing deep sequencing approach to analyze the fiber transcriptome of two allotetraploid species \( G. \) hirsutum and \( G. \) barbadense [20]. From an evolutionary point of view, these observations are surprising since the genes responsible for improved fiber properties evolved in the diploid AA genome before polyploidization [21]. None of the DD genome species produce spinnable fibers [22].

There are discrepancies in the literature regarding homeolog bias in contribution to fiber traits. Using a core set of 111 RFLP markers, Ulloa and coauthors revealed that the \( A_T \) subgenome exhibited 68% of QTLs from the five chromosomes, whereas the \( D_T \) subgenome exhibited only 32% of QTLs from the three chromosomes [23]. Another study utilizing combinations of markers found more fiber trait QTLs in the \( A_T \) subgenome than in the \( D_T \) subgenome [24]. The expression analysis of ESTs derived from immature ovules of \( G. \) hirsutum TM-1 revealed significant enrichment in all functional categories for \( A_T \) subgenome ESTs [25].
Figure 3. RNA-seq and rt-qPCR analysis detected reciprocal expression biases as a result of mutation. Original RNA-seq data are shown on left. Asterisks indicate significant (p-value <0.05) difference in gene expression level between A1 vs. D1 subgenomes in wild type (black) and mutant (blue) developing fibers. Error bars represent standard deviation from two biological replicates for RNA-seq data and three biological replicates for RT-qPCR.

doi:10.1371/journal.pone.0090830.g003
These inconsistencies could be explained by technical limitations. The QTL studies reported in the current literature are detecting only a small subset of the genes related to fiber traits that may not cover the whole genome and could be insufficient to conclude which subgenome more significantly contributes to fiber properties [19]. The microarray studies evaluated a limited number of homeologous gene pairs, resulting in limited statistical power [16,26]. Lacape and coauthors used next generation DNA sequencing technology for fiber transcriptome analysis; however, they evaluated only 617,000 good quality reads from four libraries without biological replication [20]. Unlike previous studies we obtained ~160 million reads per sample for each of two biological replicates (Table 1), providing ~5.6 times coverage of the G. hirsutum genome (~2.83 Gb per haploid [27]), which is more than enough to deliver statistically powerful transcriptional analysis. Our observation of higher expression of A1 than D1 genes in the fiber transcriptome is consistent with the results of cotton ovules ESTs analysis [25] and reflects the evolutionary role of the AA diploid progenitor in fiber quality traits of allotetraploid cotton.

It is interesting to note that the Li2 mutation coincides with an increase in the number of expressed genes, but the homeolog expression bias was significantly decreased in Li2 fiber. How expression of homeologous genes is regulated in polyploids is still unclear, although it could involve altered regulatory interactions and rapid genetic and epigenetic changes in subgenomes [28]. The evolution homeolog-specific expression after polyploidization has been extensively studied in allotetraploid cotton. Higher rates of homeolog expression bias in natural allotetraploids than in hybrid
and synthetic polyploid cottons suggested that the extent of homeolog expression bias increases over time from hybridization through evolution [26,29,30]. The Li2 mutation is negative for desirable fiber quality traits, resulting in extremely short lint fiber. Significant reduction of homeolog expression bias in short fiber suggests that the extent of homeolog expression bias is also important for fiber quality characteristics.

We observed a reciprocal switch for some genes in expression bias between homeologs during fiber developmental stages in the mutant. A high degree of expression differences between homeologous genes that are developmentally and stress regulated was reported in cotton [17,31,32]. A high-resolution genome-specific study of expression profiling for 63 gene pairs in 24 tissues in allopolyploid and their diploid progenitor cotton species demonstrated that the majority of expression differences between homeologs are caused by cis-regulatory divergence between the diploid progenitors; however, some degree of transcriptional neofunctionalization was detected as well [32].

The Li2 mutation was mapped to the DT subgenome [6–8]; however, the mutated gene and the nature of mutation are currently unknown. The greater mutation effect on the DT than on the AT subgenome observed here suggests two possible mechanisms. The network of regulatory interactions may have been interrupted by a mutation in the DT subgenome resulting in transactivation or repression of individual gene expression levels and expression cascades. Alternatively an epigenetic modulation

| Class | Identifier | TAIR10 best hit | TAIR 10 symbol | \(A_r\) biased (wtA/wtD) | \(L/A\) wtA | \(D_r\) biased (wtD/wtA) | \(L/D\) wtD |
|-------|------------|-----------------|----------------|--------------------------|------------|--------------------------|------------|
| AP2/EREBP | Gorai.006G222600 | AT4G13620 | 3.78 | 0.35 | 0.26 | 1.47 |
| AP2/EREBP | Gorai.007G010200 | AT5G52020 | 2.22 | 0.28 | 0.45 | 0.68 |
| bHLH | Gorai.007G005700 | AT1G73830 | BEE3 | 0.14 | 1.39 | 7.36 | 0.37 |
| bHLH | Gorai.006GI115500 | AT2G27230 | LHW | 2.71 | 2.45 | 0.37 | 2.87 |
| bHLH | Gorai.007G157700 | AT3G07340 | 5.10 | 0.50 | 0.20 | 1.77 |
| bHLH | Gorai.001G277600 | AT3G26744 | ICE1,5CRM | 5.82 | 0.22 | 0.17 | 0.64 |
| bHLH | Gorai.009G219100 | AT4G33880 | RSL2 | 0.37 | 1.53 | 2.68 | 0.31 |
| C2C2(Zn) CO-like | Gorai.009G065600 | AT5G24930 | COL4 | 2.38 | 2.17 | 0.42 | 2.23 |
| C2H2(Zn) | Gorai.009G092800 | AT2G29660 | 3.56 | 0.40 | 0.28 | 0.77 |
| C3H(Zn) | Gorai.011G126400 | AT5G24930 | 31.56 | 0.44 | 0.03 | 1.88 |
| CPP(Zn) | Gorai.013G121300 | AT3G22780 | TSO1 | 0.33 | 1.27 | 0.35 | 0.38 |
| G2-like | Gorai.008G224400 | AT1G32240 | KAN2 | 0.31 | 0.93 | 3.20 | 0.41 |
| Homeobox | Gorai.003G094500 | AT3G04700 | HAT5,HB-1 | 0.27 | 0.76 | 3.76 | 0.47 |
| MYB | Gorai.006G195700 | AT4G26880 | MYB85 | 12.21 | 0.31 | 0.08 | 0.56 |
| MYB | Gorai.001G015200 | AT5G35550 | MYB123,TT2 | 3.05 | 2.13 | 0.33 | 2.08 |
| MYB | Gorai.007G037100 | AT5G62470 | MYB96 | 4.72 | 2.30 | 0.21 | 5.78 |
| WRKY | Gorai.009G066900 | AT4G13620 | WRKY11 | 6.87 | 0.44 | 0.15 | 0.77 |
| bZIP | Gorai.013G258300 | AT3G44660 | DPB2 | 0.27 | 0.78 | 3.68 | 0.37 |
| AS2, LOB | Gorai.013G016900 | AT3G11090 | LBD21 | 0.29 | 1.67 | 3.48 | 0.18 |
| Aux/IAA | Gorai.009G132300 | AT3G15540 | IAA19,MSG2 | 0.33 | 0.97 | 3.07 | 0.41 |
| Aux/IAA | Gorai.010G227800 | AT3G15540 | IAA19,MSG2 | 0.10 | 0.20 | 9.65 | 0.21 |
| Aux/IAA | Gorai.001G242900 | AT3G2050 | AXR2,IAA7 | 0.24 | 0.15 | 4.23 | 0.06 |
| Aux/IAA | Gorai.006G24600 | AT4G14550 | IAA14,5LR | 4.06 | 0.28 | 0.24 | 0.43 |
| Aux/IAA | Gorai.001G043900 | AT4G23280 | IAA29 | 2.39 | 2.20 | 0.42 | 2.60 |
| Aux/IAA | Gorai.007G277000 | AT5G3700 | IAA4 | 0.19 | 0.32 | 5.28 | 0.25 |
| General | Gorai.011G240900 | AT3G52270 | | 3.94 | 0.40 | 0.25 | 0.40 |
| GATA type (Zn) | Gorai.005G230900 | AT1G0200 | WLM1 | 3.86 | 0.30 | 0.26 | 0.62 |
| TAZ domain | Gorai.007G056400 | AT5G67480 | BT4 | 0.09 | 0.99 | 11.08 | 0.19 |
| GRP | Gorai.008G029100 | AT4G39260 | GRP8 | 3.23 | 0.47 | 0.31 | 0.74 |
| unclassified | Gorai.003G092990 | AT2G01818 | | 3.23 | 0.26 | 0.31 | 0.39 |
| unclassified | Gorai.013G258400 | AT2G27580 | | 2.85 | 0.38 | 0.35 | 0.82 |
| unclassified | Gorai.012G146800 | AT3G18870 | | 0.47 | 1.31 | 2.14 | 0.45 |
| unclassified | Gorai.008G221300 | AT4G19050 | | 2.08 | 2.23 | 0.48 | 2.01 |
| unclassified | Gorai.010G120300 | AT5G0770 | | 0.06 | 1.17 | 17.63 | 0.36 |
| unclassified | Gorai.004G118800 | AT2G3750 | | 2.79 | 0.48 | 0.36 | 0.47 |

The genes significantly (FDR corrected \(p\)-value < 0.05) down-regulated or up-regulated more than 2-fold are shown in boldface and underlined. doi:10.1371/journal.pone.0090830.t003

PLOS ONE | www.plosone.org 7 March 2014 | Volume 9 | Issue 3 | e90830
may preferentially target the DT subgenome. It has been shown that small RNAs can control gene expression and epigenetic regulation in response to hybridization [33–36]. For example, miRNAs in allopolyploid Arabidopsis triggered unequal degradation of parental target genes [33]. Similarly, in rice hybrids small RNA populations inherited from parents were responsible for biased expression [36]. Additional investigations of epigenetic and chromatin level modifications will provide insights into causes of gene expression variation between subgenomes.

| Description | Identifier | TAIR10 best hit | \(A_r\) biased (wtA/wtD) | \(L_{i2}A/wtA\) | \(D_r\) biased wtD/wtA | \(L_{i2}D/wtD\) |
|-------------|------------|----------------|------------------|----------------|------------------|----------------|
| Hemicellulose synthesis | Goral.011G272700 | AT3G45400 | 4.82 | 1.16 | 0.21 | 2.03 |
| Xyloglucan \(\beta\)-galactosyltransferase; MUR3 | Goral.007G150100 | AT2G20370 | 3.36 | 0.41 | 0.30 | 0.97 |
| Xylosyltransferase; IRX9 | Goral.006G168500 | AT2G37090 | 0.28 | 0.64 | 3.63 | 0.42 |
| Cell wall proteins | Fasciclin-like arabinogalactan 7 | Goral.001G219000 | AT2G04870 | 9.65 | 0.41 | 0.10 | 1.72 |
| Arabinogalactan protein 16 | Goral.007G023300 | AT2G46330 | 3.07 | 0.42 | 0.33 | 0.44 |
| Fasciclin-like arabinogalactan 10 | Goral.003G132100 | AT3G46550 | 0.05 | 0.25 | 20.68 | 0.43 |
| Fasciclin-like arabinogalactan 10 | Goral.007G092300 | AT3G60900 | 0.09 | 2.30 | 11.31 | 2.30 |
| Arabinogalactan protein 18 | Goral.013G048900 | AT4G37450 | 3.01 | 0.91 | 0.33 | 2.51 |
| Fasciclin-like arabinogalactan 17 | Goral.006G147500 | AT5G06390 | 2.08 | 0.42 | 0.48 | 0.59 |
| Fasciclin-like arabinogalactan 1 | Goral.013G152900 | AT5G55730 | 3.71 | 2.62 | 0.27 | 1.65 |
| Proline-rich protein 2 | Goral.002G193500 | AT2G21140 | 0.23 | 0.89 | 4.35 | 0.37 |
| Reversibly glycosylated polypeptide 1 | Goral.001G090200 | AT3G02230 | 0.33 | 2.04 | 3.07 | 1.99 |
| Degradation | Peptidoglycan-binding LysM | Goral.008G281400 | AT5G62150 | 19.29 | 0.35 | 0.05 | 1.68 |
| Glycosyl hydrolase 9B8 | Goral.007G170600 | AT2G32990 | 0.10 | 1.57 | 10.48 | 0.37 |
| Glycosyl hydrolase | Goral.006G118100 | AT5G20950 | 0.17 | 1.33 | 6.02 | 0.49 |
| \(\alpha\)-L-fucosidase 1 | Goral.008G282700 | AT2G28100 | 4.69 | 0.86 | 0.21 | 2.36 |
| \(\beta\)-Xylosidase 1 | Goral.011G198200 | AT5G49360 | 2.01 | 1.88 | 0.50 | 2.25 |
| \(\beta\)-Xylosidase 2 | Goral.008G140600 | AT1G02640 | 0.31 | 0.71 | 3.18 | 0.22 |
| Rhamnogalacturonate lyase | Goral.002G138200 | AT1G09890 | 3.63 | 1.60 | 0.28 | 3.51 |
| Rhamnogalacturonate lyase | Goral.007G231600 | AT2G22620 | 0.40 | 2.06 | 2.53 | 2.33 |
| Rhamnogalacturonate lyase | Goral.008G093300 | AT2G22620 | 11.16 | 0.20 | 0.09 | 6.32 |
| Polygalacturonase 2 | Goral.002G161400 | AT1G07370 | 0.35 | 0.46 | 2.85 | 0.38 |
| Polygalacturonase 2 | Goral.005G136300 | AT1G07370 | 5.03 | 0.39 | 0.20 | 1.55 |
| Pectin lyase-like | Goral.004G139900 | AT4G23500 | 0.36 | 1.48 | 2.81 | 2.50 |
| Pectin methylesterase 1 | Goral.009G147800 | AT1G38400 | 11.39 | 2.27 | 0.09 | 3.18 |
| Pectinacetylesterase | Goral.004G176200 | AT4G19410 | 0.38 | 2.11 | 2.66 | 1.89 |

The genes significantly (FDR corrected \(p\)-value <0.05) down-regulated or up- regulated more than 2-fold are shown in boldface and underlined.

doi:10.1371/journal.pone.0090830.t004

TRs and Cell Wall Functional Categories of Genes Regulated by \(L_{i2}\) Mutation

Previous transcriptomics and metabolomics studies have shown that the \(L_{i2}\) mutation terminated the cotton fiber elongation process [6,10]; therefore, genes with changed expression level in the mutant could be involved in elongation. In the present work, we described in detail TRs and cell wall functional categories, which are critical for fiber developmental processes. Many genes
in this list (Table 3 and Table 4) were not functionally characterized in cotton; although, based on sequence similarity to genes characterized in Arabidopsis, they could be involved with fiber elongation and represent candidates for further functional analysis in cotton.

**Transcriptional regulators.** Many TRs regulated by Li2 mutation are involved in hormonal signaling and development. Particularly, two AP2/EREBPs and six Aux/IAAs were in the pool of TRs affected by Li2 mutation (Table 3). Plant hormones are important for fiber development. It is well documented that exogenous applications of auxins and gibberellic acid stimulate the differentiation of fibers and promote elongation, while abscisic acid and cytokinins inhibit fiber growth in an in vitro cotton ovule culture system [37,38]. Among the auxin responsive genes, Gorai.009G132300 and Gorai.010G227800, whose transcript abundances were significantly reduced in the D1 genome of Li2, showed sequence similarity to grapevine VvIAA19 regulator [39]. Transgenic Arabidopsis plants over expressing VvIAA19 exhibited faster growth, including root elongation and floral transition, than the control, suggesting that grape Aux/IAA19 protein is likely to play a crucial role as a plant growth regulator. In the group of bHLH family of TRs, Gorai.007G057500 transcript abundance was significantly reduced in the D1 genome of Li2 and showed sequence similarity to Arabidopsis BEE3, one of several redundant positive regulators of brassinosteroids signaling required for normal growth and development [40].

The actin cytoskeleton plays an important role in cell morphogenesis; down-regulation of GlbACT1 disrupted the actin cytoskeleton network in fibers that resulted in inhibition of fiber elongation [41]. A GATA type TR Gorai.005G2230900, a homolog of Arabidopsis WLM1, was down-regulated in the D1 subgenome of Li2; a recent study revealed that plant LIM-domain containing proteins (LIMs) define a highly specialized actin binding protein family, which contributes to the regulation of actin bundle in virtually all plant cells [42].

**Cell wall.** The plant cell wall has a dual role during elongation: to sustain the large mechanical forces caused by cell turgor and to permit controlled polymer extension generating more space for protoplast enlargement [43]. The active biosynthesis of matrix polysaccharides along with increased activity of cell wall loosening enzymes has been considered to be associated with cell wall extension [44–46]. Expression levels of genes encoding enzymes involved in xyloglucan and glucuronoxylan biosynthesis were decreased as a result of Li2 mutation. Particularly, xyloglucan β-galactosyltransferase (Arabidopsis homolog, MUR3 [49]) and xylosyltransferase (IRX9 [30]) were down-regulated in the A1 or D1 subgenomes of mutant fibers (Table 4).

Among cell wall proteins arabinoxylans were the most abundant members. Arabinoxylan-proteins have been implicated in many processes involved in plant growth and development, including cell expansion [51,52].

Primary cell wall expansibility and strength is in part mediated by a group of enzymes that comprise a large family of cell wall modifying proteins, the xyloglucan endotransglycosylase/hydrolases (XTHs). XTHs are apoplast-localized enzymes that cleave and react with xyloglucan polymers [53,54]. The role of XTHs in cotton fiber elongation has been demonstrated: transgenic overexpression of GlbXTH1 in cotton increased fiber length up to 20% [55]. D1 biased Gorai.007G057400 corresponding to GlbXTH1 was down-regulated in mutant fiber.

**Conclusion**

Repeated polyploidization over evolutionary time has played a significant role in adding genetic variation to the genomes of plant species. The evolution of the homeolog expression after polyploidization has been extensively studied in cotton comparing expression profiling between parental diploids and natural and synthetic allotetraploid species. This is the first report that explored the effects of a single mutation on the homeolog expression of allotetraploid cotton. Our results showed that significant reduction of the homeolog expression bias in mutant fiber correlates with negative fiber traits, indicating that the extent of homeolog expression bias is important for fiber quality characteristics. In addition, we observed significantly greater mutation effects on the D1 than on the A1 subgenome that might be explained by localization of the mutated gene. Additional studies using numerous naturally occurring cotton fiber mutations are needed to confirm these observations. This work will lead to an understanding of how gene regulation between A1 and D1 homeologs contributes to enhanced fiber morphology in cultivated cotton allotetraploids.

**Materials and Methods**

**Plant Material and RNA Isolation**

The cotton short fiber mutant Li2 was developed as a near-isogenic line (NIL) with the WT upland cotton line DP5690 as described before [6]. Growth conditions and fiber sampling were previously described [6]. Cotton bolls were harvested at the following time-points during development: day of anthesis (DOA), 1, 3, 5, 8, 12, 16, and 20 days post-anthesis (DPA). Cotton fibers were isolated from developing ovules using a glass bead shearing technique to separate fibers from the ovules [56]. Total RNA was isolated from detached fibers using the Sigma Spectrum Plant Total RNA Kit (Sigma-Aldrich, St. Louis, MO) with the optional on column DNase1 digestion according to the manufactuer's protocol. The concentration of each RNA sample was determined using a NanoDrop 2000 spectrophotometer (NanoDrop Technologies Inc., Wilmington, DE). The RNA quality for each sample was determined by RNA integrity number (RIN) using an Agilent Bioanalyzer 2100 and the RNA 6000 Nano Kit Chip (Agilent Technologies Inc., Santa Clara, CA) with 250 ng of total RNA per sample.

**RT-qPCR Analysis**

The experimental procedures and data analysis related to RT-qPCR were performed according to the Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) guidelines [57]. Eight fiber developmental time-points mentioned above were used for RT-qPCR analyses of homeolog pairs which showed reciprocal expression biases. Only one time point, 8 DPA, was used for RT-qPCR confirmation of RNA-seq data of selected genes. The detailed description of reverse transcription, qPCR and calculation were previously reported [6]. Single nucleotide polymorphisms that distinguish the A1 and D1 subgenome copies of the selected genes were identified by aligning reads from the RNA-seq data to the *G. raimondii* reference mRNAS sequences [11]. These homeologous SNPs were used to design subgenome specific primers by the SNAPER approach, whereby an additional mismatch is included near the end of the SNP-specific primers to increase stringency [58]. Primer sequences are provided in Table S3 and Table S4 in File S1. Correlations of biased expression patterns between RNA-seq and RT-qPCR data were calculated using GraphPad Prism 5 software (Pearson test).
Library Preparation and Sequencing

RNA samples from L2 and wild type cotton fiber at 8 DPA (in two biological replicates) were subjected to paired-end Illumina mRNA sequencing (RNA-seq). Library preparation and sequencing were conducted by Data2Bio LLC (2079 Roy J. Carver Laboratory, Ames, Iowa). Indexed libraries were prepared using the Illumina protocol outlined in the TrueSeq RNA Sample Prep Guide (Part # 15000136 Rev. A, November 2010). The library size and concentration were determined using an Agilent Bioanalyzer. The indexed libraries were combined and seeded onto one lane of the flowcell. The libraries were sequenced using 101 cycles of chemistry and imaging, resulting in paired-end (PE) sequencing reads with length of 2 × 101 bp. The raw reads were submitted to the Sequence Read Archive (accession number SRP026301).

Processing of Illumina RNA-Seq Reads and Mapping to AT and D1 Subgenomes of Gossypium hirsutum

The reads were trimmed with SICKLE (https://github.com/najoshi/sickle) using a quality score cutoff of 20. Mapping the reads (in pairs where both reads of a pair passed trimming) to the 13 chromosomes of the G. raimondii genome D1 v2 reference sequence was performed using GSNAP [59]. Default parameters were used, but with the flags “-n 1 -Q” which means that only a single mapping was reported for each read, and reads with multiple equally good hits were thrown away rather than randomly mapped. We used a cotton SNP index generated between DD genome G. raimondii and the AA genome G. arboreum to categorize reads of the allotetraploid G. hirsutum as belonging to the AT or D1 subgenomes according to the method reported previously [14].

Digital Gene Expression Analysis

The comparison of the number of reads mapped to the genes of G. raimondii reference genome was used as an indicator of the relative digital gene expression (DGE). The JMP/Genomics 6.0 (SAS, Cary, NC, USA) was used for data normalization and statistical analysis. The data was normalized using TMM (Trimmed Mean of M component) method [60]. Genes with less than 10 reads in one sample were removed before normalization; from 37,223 genes assigned to chromosomes, 31,114 genes passed filtering conditions and were processed for normalization. The ANOVA process was fit to the normalized data, with the data deviation from two biological replicates for RNA-seq data and three biological replicates for RT-qPCR. Bar charts represent RNA-seq and RT-qPCR data (side by side) at 8 DPA of fiber development for 8 randomly selected genes from Table S1 in File S1. Error bars indicate standard deviation from two biological replicates for RNA-seq data and three biological replicates for RT-qPCR.

Supporting Information

Figure S1 RT-qPCR confirmation of silencing or activation of genes as a result of mutation. Bar charts represent RNA-seq and RT-qPCR data (side by side) at 8 DPA of fiber development for 11 randomly selected genes from Table 3 and Table 4. Pearson correlation (GraphPad Prism 5 software) of expression patterns for selected genes between RNA-seq and RT-qPCR data is provided in the table; correlation coefficients with p-value less than 0.05 are shown in boldface and underlined. Error bars indicate standard deviation from two biological replicates for RNA-seq data and three biological replicates for RT-qPCR.

File S1 Supporting tables. Table S1. Silencing or activation of genes as a result of mutation. Table S2. Mutation effects on functional distribution of homeolog genes. Fisher’s exact test results. Table S3. Primer’s sequences for detection expression of homeolog pairs. Table S4. Primer’s sequences.

Acknowledgments

We thank Data2Bio for sequencing service. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture that is an equal opportunity provider and employer.

Author Contributions

Conceived and designed the experiments: MN DDF. Performed the experiments: MN DJH CF. Analyzed the data: MN GT KMY JTP JAU. Contributed reagents/materials/analysis tools: DJH JTP JAU. Wrote the paper: MN GT DDF DJH JTP JAU. Read and approved the final manuscript: MN GT DDF DJH CF KMY JTP JAU.

References

1. Wendel JF, Cronn RC (2002) Polyploidy and the evolutionary history of cotton. Advances in Agronomy 78: 139–196.
2. Wendel JF (1989) New World tetraploid cottons contain Old World cytoplasm. Proceedings of the National Academy of Sciences of the United States of America 86: 4132–4136.
3. Bara AS, Malik AG (1984) Development of the cotton fiber. International Review of Cytology 89: 65–113.
4. Kim HJ, Triplett BA (2001) Cotton fiber growth in planta and in vitro. Models for plant cell elongation and cell wall biogenesis. Plant physiology 127: 1361–1366.
5. Narbath EV, Kohel RJ (1990) Inheritance and Linkage Analysis of a New Fiber Mutant in Cotton. Journal of Heredity 81: 131–135.

6. Hinchliffe DJ, Turley RB, Naoumkina M, Kim HJ, Tang Y, et al. (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.

7. Kohel RJ, Stelly DM, Yu J (2002) Tests of six cotton (Gossypium hirsutum L.) mutants for association with aneuploidy. The Journal of Heredity 93: 130–132.

8. Page JT, Gingle AR, Udall JA (2013) PolyCat: a resource for genome categorization of sequencing reads from allopolyploid organisms. G3: Genes|Genomes|Genetics 14: 153.

9. Page JT, Gingle AR, Udall JA (2013) PolyCat: a resource for genome categorization of sequencing reads from allopolyploid organisms. G3: Genes|Genomes|Genetics 14: 153.

10. Dong S, Adams KL (2011) Differential contributions to the transcriptome of allotetraploid cotton and its diploid progenitors. The New phytologist 190: 1045–1057.

11. Hovav R, Udall JA, Chaudhary B, Rapp R, Flagel L, et al. (2008) Partitioned expression level dominance in allopolyploid cotton. Heredity 110: 171–180.

12. Chaudhary B, Flagel L, Stupar RM, Udall JA, Verma N, et al. (2009) Reciprocal silencing, transcriptional bias and functional divergence of homologs in polyploid cotton (Gossypium). Genetics 182: 503–517.

13. Ha M, Lu J, Tian L, Ramachandran V, Kasschau KD, et al. (2009) Small RNAs serve as a genetic buffer against the genomic shock in Arabidopsis interspecific hybrids and allopolyploids. Proceedings of the National Academy of Sciences of the United States of America 106: 17835–17840.

14. Madlung A, Tyagi AP, Watson B, Jiang H, Koguchi T, et al. (2005) Genomic changes in synthetic Arabidopsis polyploids. The Plant journal 41: 221–230.

15. Salomon A, Ainsouche M, Wendel JF (2005) Genetic and epigenetic consequences of recent hybridization and polyplody in Spartina (Poaceae). Molecular Ecology 14: 1163–1173.

16. He G, Zhu X, Eling AA, Chen L, Wang X, et al. (2010) Global epigenetic and transcriptional trends among two rice subspecies and their reciprocal hybrids. The Plant cell 22: 17–33.

17. Beasley CA, Ting IT (1973) The effects of plant growth substances on in-vitro fiber development from fertilized cotton ovules. American Journal of Botany 60: 130–139.

18. Beasley CA, Ting IT (1974) Effects of plant growth substances on in-vitro fiber development from unfertilized cotton ovules. American Journal of Botany 61: 180–194.

19. Kolno M, Takano H, Horisuchi H, Fujita K, Suzaki S (2012) Auxin-nonresponsive grape Aux/IAA19 is a positive regulator of plant growth. Molecular biology reports 39: 911–917.

20. Friedrichs DM, Nemhauser J, Marrutis T, Mallof JN, Alonso J, et al. (2002) Three redundant brassinosteroid early response genes encode putative bHLH transcription factors required for normal growth. Genetics 162: 1445–1456.

21. Li XB, Fan XP, Wang XL, Cai L, Yang WC (2005) The Cotton ACTIN1 gene is functionally expressed in fibers and participates in fiber elongation. The Plant cell 17: 859–875.

22. Papuga J, Hoffmann C, Dieteler M, Moes D, Moraar F, et al. (2010) Arabidopsis LIM proteins: a family of actin bundles with distinct expression patterns and modes of regulation. The Plant cell 22: 3034–3052.

23. Cosgrove DJ (2001) Wall structure and wall loosening. A look backwards and forwards. Plant physiology 125: 131–134.

24. Takeda T, Furuta Y, Aoki T, Murano M, Kuzushi Y, et al. (2002) Suppression and acceleration of cell elongation by integration of xyloglucan in pea stem segments. Proceedings of the National Academy of Sciences of the United States of America 99: 9053–9060.

25. Park KY, Baba K, Furuta Y, Ida S, Sainoumika K, et al. (2004) Enhancement of growth and cellulose accumulation by overexpression of xyloglucanase in poplar. FEBS letters 543: 161–167.

26. Hayashi T, Delmer DP (1988) Xyloglucan in the cell walls of cotton. Carbohydrate research 181: 273–277.

27. Howdyler HR, Franz G, Meier H (1979) Changes in the composition of cotton fiber cell walls during development. Planta 146: 635–642.

28. Shimizu Y, Aotsuka S, Hasegawa O, Kawada T, Sakamo T, et al. (1997) Changes in levels of mRNAs for cell wall-related enzymes in growing cotton fiber cells. Plant & cell physiology 38: 373–378.

29. Madison M, Dunand C, Li X, Verma R, Vanzin GF, et al. (2003) The MUR3 gene of Arabidopsis encodes a xyloglucan galactosyltransferase that is evolutionarily related to animal exostosins. The Plant cell 15: 1662–1670.

30. Lee C, Zhong R, Ye ZH, Zhang K, Wang X, et al. (2009) Three redundant brassinosteroid early response genes are required for elongation of the xylan backbone. Plant & cell physiology 53: 135–143.

31. Lee JK, Sakata Y, Mau SL, Petridino F, Baccic A, et al. (2005) Arabidopsis galectin proteins are required for appressorium cell extension in the moss Physcomitrella patens. The Plant cell 17: 3051–3063.

32. Yang J, Showalter AM (2007) Expression and localization of AraAGP18, a lysine-rich arabinogalactan-protein in Arabidopsis. Planta 226: 169–179.

33. Nishitani K, Tomimasa R (1992) Endo-xylanase-transgenics, a novel class of glycosyltransferase that catalyzes transfer of a segment of xyloglucan molecule to another xyloglucan molecule. The Journal of biological chemistry 267: 21058–21064.

34. Fry SC, Smith RC, Renwick KF, Martin DJ, Hodge SK, et al. (1992) Xyloglucan endotransglucosylase, a new wall-loosening enzyme activity from plants. The Biochemical Journal 282 (Pt 3): 821–828.

35. Lee J, Burns TH, Light G, Sun Y, Fokar M, et al. (2010) Xyloglucan endotransglucosylase/hydrolase genes in cotton and their role in fiber elongation. Planta 232: 1190–1207.

36. Talerico EW, Beckin D (2007) Analysis of gene expression in cotton fiber initials. BMC plant biology 7: 22.

37. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, et al. (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clinical chemistry 55: 611–622.

38. Drenkard E, Richter BG, Rozen S, Studits LM, Angell NA, et al. (2000) A simple procedure for the analysis of single nucleotide polymorphisms facilitates map-based cloning in Arabidopsis. Plant physiology 124: 1483–1492.

39. Wu TD, Nacu S (2010) Fast and SNP-tolerant detection of complex variants and splicing in short reads. Bioinformatics 26: 873–881.

40. Robinson MD, Oshlack A (2010) A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol 11: R25.
61. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B (Methodological) 57: 289–300.

62. Thimm O, Blasing O, Gibon Y, Nagel A, Meyer S, et al. (2004) MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. The Plant journal 37: 914–939.