Phylogenetically Distinct Cellulose Synthase Genes Support Secondary Wall Thickening in Arabidopsis Shoot Trichomes and Cotton Fiber

Lissete Betancur[1], Bir Singh[2], Ryan A. Rapp[3], Jonathan F. Wendel[3], M. David Marks[4], Alison W. Roberts[5] and Candace H. Haigler[1,2*]

[1]Department of Plant Biology, North Carolina State University, Box 7612, Raleigh NC 27695-7612, USA

[2]Department of Crop Science, North Carolina State University, Box 7620, Raleigh NC 27695-7620, USA

[3]EEOB Department, Iowa State University, 251 Bessey Hall, Ames, IA 50011, USA

[4]Department of Plant Biology, University of Minnesota, 1445 Gortner Ave, St. Paul, MN 55108, USA

[5]Department of Biological Sciences, CBLS, University of Rhode Island, Kingston, RI 02881, USA

Current Address: Ryan A. Rapp, Department of Molecular Breeding, Monsanto Company, 800 North Lindbergh Blvd., Mail Stop N3SA, St. Louis, MO 63167, USA

[*]Corresponding author

Tel: +1 919 515 5645; Fax: +1 919 515 5315; E-mail: candace_haigler@ncsu.edu

Available online on 5 February 2010 at www.jipb.net and www.interscience.wiley.com/journal/jipb

doi: 10.1111/j.1744-7909.2010.00934.x

Abstract

Through exploring potential analogies between cotton seed trichomes (or cotton fiber) and arabidopsis shoot trichomes we discovered that CesAs from either the primary or secondary wall phylogenetic clades can support secondary wall thickening. *CesA* genes that typically support primary wall synthesis, *AtCesA1,2,3,5,* and *6*, underpin expansion and secondary wall thickening of arabidopsis shoot trichomes. In contrast, apparent orthologs of *CesA* genes that support secondary wall synthesis in arabidopsis xylem, *AtCesA4,7*, and *8*, are up-regulated for cotton fiber secondary wall deposition. These conclusions arose from: (a) analyzing the expression of *CesA* genes in arabidopsis shoot trichomes; (b) observing birefringent secondary walls in arabidopsis shoot trichomes with mutations in *AtCesA4, 7,* or *8*; (c) assaying up-regulated genes during different stages of cotton fiber development; and (d) comparing genes that were co-expressed with primary or secondary wall CesAs in arabidopsis with genes up-regulated in arabidopsis trichomes, arabidopsis secondary xylem, or cotton fiber during primary or secondary wall deposition. Cumulatively, the data show that: (a) the xylem of arabidopsis provides the best model for secondary wall cellulose synthesis in cotton fiber; and (b) *CesA* genes within a "cell wall toolbox" are used in diverse ways for the construction of particular specialized cell walls.

Introduction

The aim of this research was to compare gene expression patterns associated with secondary cell wall synthesis in shoot trichomes of arabidopsis (*Arabidopsis thaliana*) and cotton (*Gossypium hirsutum*) fibers, which are unicellular seed epidermal trichomes. Thick secondary walls are deposited after cell expansion to confer extra strength to some specialized plant cells, and their industrial uses range from textile and wood pulp fibers to forage and biofuel feedstocks. Cotton fibers provide humankind's most important renewable textile fiber. They differentiate over a period of ∼45–60 d, beginning with expansion of individual cells above the plane of the ovule outer integument near the day of flower opening. The

temporal progression of cotton fiber development is described by days post anthesis (DPA). The differentiation program includes distinct phases of primary and secondary wall deposition to support fiber elongation and cell wall thickening, respectively. The >90% cellulose content in the secondary wall and the >2.25 cm fiber length of domesticated cotton fiber are key characters that have made this unusual seed epidermal hair especially useful to humans for over 7 000 years (Ryser 1985; Dillehay et al. 2007).

Because cotton fibers are a type of trichome, researchers have frequently looked for analogies with vegetative trichomes, especially those in the model plant arabidopsis. An alternative model cell type for investigating the function of genes expressed in cotton fiber would be particularly useful given the substantial time and resources required for stable cotton transformation. The term trichom (German) / trichome (English) was derived in 1875 from the Greek trikhōma, meaning 'growth of hair', to denote a filamentous outgrowth such as a plant epidermal hair (Merriam-Webster's Collegiate Dictionary, 11th Ed. 2003). 'Trichome' now refers to outgrowths of the plant epidermis exclusive of cells of sub-epidermal origin, mainly leaf and stem (shoot) trichomes, cotton seed hairs (fiber), and root hairs. These morphologically diverse cells and multicellular structures have been shaped by natural selection for a wide range of functions, including water absorption, thermoregulation, UV protection, both repellant and attractant components of defense, and seed dispersal (Evert 2006). The seed epidermal trichomes that presently dominate world cotton commerce were derived through domestication of wild cotton species with ancestral fibers that probably evolved to aid seed dispersal and/or germination (through retaining water around the seed) (Wendel et al. 2009).

Given this diversity, it is of interest to explore the homologies among the various epidermal outgrowths classified as trichomes. Unicellular, apparently unbranched, spiny epidermal projections are present in the fossils of the earliest land plants, e.g. *Trichopherophyton teuchansii* (Lyon and Edwards 1991). Nonetheless, the various types of epidermal outgrowths in extant plants likely evolved independently and on multiple occasions rather than once from a common ancestor. This is reflected in the many origins of trichomes on different plant organs (shoots, roots, sepals, seeds, etc.) and the many differences in the cellular characteristics of trichomes. For example, the unicellular and branched leaf trichomes of arabidopsis arise through a different regulatory network than the multicellular trichomes of snapdragon (*Antirrhinum majus*) and solanaceous species such as tobacco (Serna and Martin 2006; Jakoby et al. 2008).

Although trichomes are clearly "convergent" at the gross structural level, it also is clear that true parallelisms exist among structures across species. For example, cross-species gene over-expression and complementation experiments have identified common transcription factors involved in the initiation of arabidopsis leaf trichomes and cotton fiber (reviewed in Lee et al. 2007; Guan et al. 2007; Liu and Zhang 2008; Boopathi and Ravikesavan 2009). This finding is consistent with the close phylogenetic relationship of arabidopsis (Brassicales) and cotton (Malvales) and the common character of unicellularity in the two types of trichomes. Mapping studies in cotton have also shown close linkage or common genetic control of leaf trichomes and cotton fiber (Desai et al. 2008 and references therein). Similarly, the process of cell elongation that is dependent on cytoskeletal activity, generation of turgor pressure, and primary wall synthesis is likely to share some common controls in shoot trichomes and cotton fiber (Jakoby et al. 2008). Similar investigations were lacking for the cell wall thickening phase of trichome differentiation. Consequently, it was not previously clear if arabidopsis trichomes were a good model for the later stages of cotton fiber development.

Given the dominance of cellulose in cotton fiber, the ten-member cellulose synthase (*CesA*) gene family in arabidopsis (Somerville 2006) provides a useful reference point for comparing cell wall biosynthetic processes between cotton fibers and arabidopsis shoot trichomes. Although their precise role in the biochemical pathway of cellulose synthesis is still undefined, the cellulose synthases are UDP-glucose: 1,4-ß-D-glucosyltransferase enzymes in the glycosyltransferase family 2 (reviewed in Somerville 2006). In arabidopsis, members of a triplet of CesA isoforms (AtCesA1, 3, 6, or a 6-like protein) have non-redundant roles in primary wall cellulose synthesis (Persson et al. 2007; Desprez et al. 2007), whereas members of another triplet (AtCesA4, 7, 8) have non-redundant roles in secondary wall cellulose synthesis in xylem cells (Taylor et al. 2003) For convenience, the six clades that arabidopsis CesAs form with their orthologs from other seed plants have been designated P1, P2, and P3 (defined by *AtCesA1, 3,* and *6,* respectively) and S1, S2, and S3 (defined by *AtCesA 4, 7,* and *8*, respectively) (Haigler and Roberts 2009). The remainder of the ten arabidopsis *CesA* genes appear to be relatively recently derived within P1 (*AtCesA10*) and P3 (*AtCesA2, 5,* and *9*) in the crucifer lineage. The arabidopsis members of P3 are functionally interchangeable (Persson et al. 2007; Desprez et al. 2007) and *AtCesA10* has limited expression (Doblin et al. 2002). Based on available evidence, the division of function between these traditionally defined primary- and secondary wall CesAs is broadly conserved in other angiosperms even if the number of *CesA* genes has increased (Tanaka et al. 2003; Djerbi et al. 2005; Kumar et al. 2009). Other genes/proteins may be required for both primary and secondary wall cellulose biosynthesis, e.g. an endo-glucanase–*Korrigan*, or alternatively may parallel the *CesA* family in having isoforms specialized for primary or secondary wall biosynthesis, e.g. within the *COBRA* and *POM/CTL* families (**Table 1**; see the review of Somerville 2006 for additional information).

Against this backdrop, we set out to address whether secondary wall cellulose synthesis is under the same or different genetic control in cotton fiber as compared to arabidopsis shoot trichomes. We used several approaches to address this question: (a) assaying the expression of nine *CesA promoter:GUS* reporter genes in arabidopsis seedlings and trichomes; (b) documenting the occurrence of birefringent thick walls in trichomes of secondary wall *CesA* mutants of arabidopsis; (c) quantifying *CesA* expression levels in isolated arabidopsis trichomes and processed shoots; (d) analyzing genes up-regulated for cotton fiber primary and secondary wall deposition; and (e) comparing the intersection between several gene expression sets from arabidopsis and cotton. The latter comparisons included data from arabidopsis secondary xylem because cotton fiber expresses apparently orthologous genes to support fiber thickening (Haigler et al. 2005; Ko et al. 2006; Haigler et al. 2009).

Cumulatively the data show that phylogenetically distinct *CesA* genes support secondary wall cellulose synthesis in arabidopsis shoot trichomes and cotton fiber. Whereas arabidopsis shoot trichomes express *CesA* genes belonging to the three primary wall clades (P1–3) throughout their development, cotton fiber up-regulates *CesA* genes belonging to the three secondary wall clades (S1–3) at the onset of wall thickening. Rather than shoot trichomes, arabidopsis xylem (and similar cortical fiber cells) is the preferred model for the secondary wall stage of cotton fiber development. More generally, the data illustrate the flexible deployment of CesA genes/proteins in either P- or S- clades to accomplish secondary wall thickening in different types of trichomes, thereby blurring the distinction previously made between primary wall-related and secondary wall-related CesAs.

Results

Expression of 9 *CesA promoter:GUS* reporter gene constructs in arabidopsis shoot trichomes

Although expression patterns for the *ProCesA::GUS* lines have been summarized (Doblin et al. 2002) and some data have been published (Holland et al. 2000; Desprez et al. 2007; Scheible et al. 2001; Persson et al. 2007), no comprehensive analysis of these lines was available and expression in trichomes had not been reported. We observed *GUS* expression for nine of ten *ProCesA::GUS* lines (**Figures 1**, **2**), but not the weak expression of *ProCesA10::GUS* in seeds reported previously (M. Doblin, personal communication), even after overnight color development. For most lines and observations, 3 h was an optimal time for color development.

As expected given their roles in primary wall synthesis, the cloned 5′ fragments of *AtCesA(1,2,3,5,6,9)* drove *GUS* expression in expanding tissues of young plants (typically

15 d post germination; **Figure 1A-F**). (See the methods section for the length of the native *CesA* promoter regions used in the reporter gene constructs.) *GUS* expression was widely distributed and strongest for the *ProCesA(1,3)::GUS* lines, which showed complete color development after 1 h. A widely distributed but intermediate level of expression was observed for the *ProCesA(2,6)::GUS* lines (which required 3 h to show strong blue color). The weakest expression was observed for *ProCesA(5,9)::GUS* lines, which did not show any blue color at 1 h (data not shown). After 3 h, the expression pattern was widely distributed for the *ProCesA5::GUS* line, whereas it was restricted to young rosette leaves of the *ProCesA9::GUS* line. Strong *AtCesA(1,2,3,5,6)* promoter activity was detected in all flower parts, whereas the *AtCesA9* promoter drove GUS expression only in pollen (data not shown).

Strong GUS activity was detected in leaf trichomes with thick walls in the *ProCesA(1,2,3,5,6)::GUS* lines after 3 h incubation (**Figure 1G–K**), but no activity was detected in the *ProCesA9::GUS* line even after 24 h color development (**Figure 1L**). We examined young trichomes at the shoot apex to test whether activity of any of the promoters was confined to the wall thickening stage. However, the *AtCesA(1,2,3,5,6)* promoters were all active in the early stages of trichome elongation (**Figure 1M–Q**). The lack of thick walls in these young trichomes was confirmed by the absence of wall birefringence in polarization microscopy (**Figure 1R–V**), whereas the thick walls of mature trichomes were bright blue when photographed under identical optical conditions (**Figure 1W**). The polarization micrographs (**Figure 1R–W**) were recorded with a first-order red (530 nm) retardation plate inserted, which increased the intensity and color contrast arising from birefringent secondary walls, if present.

As expected given the proven roles of *CesA(4,7,8)* in secondary wall synthesis, their promoters drove GUS expression in xylem of leaves and roots after an optimum incubation time of 3 h (**Figure 2A–F**). Surprisingly, even thick-walled trichomes did not stain blue after 3 h color development in these lines (**Figure 2G–I**). Isolated trichomes were also observed after 24 h color development, and the results were still negative except for occasional faint blue color, especially in the *ProCesA7::GUS* line (data not shown). In addition, limited GUS activity at the top of the style and in anthers was found in reproductive parts for these lines (data not shown).

Analysis of *CesA* expression in arabidopsis trichomes and processed shoots

As reporter gene assays may misrepresent native gene expression patterns due to omitted regulatory elements in the cloned 5′ fragment of the gene, *CesA* expression in isolated trichomes and processed shoots was assayed by quantitative RT-PCR (qPCR) using gene-specific primers (**Table 2**; **Figure 3A**).
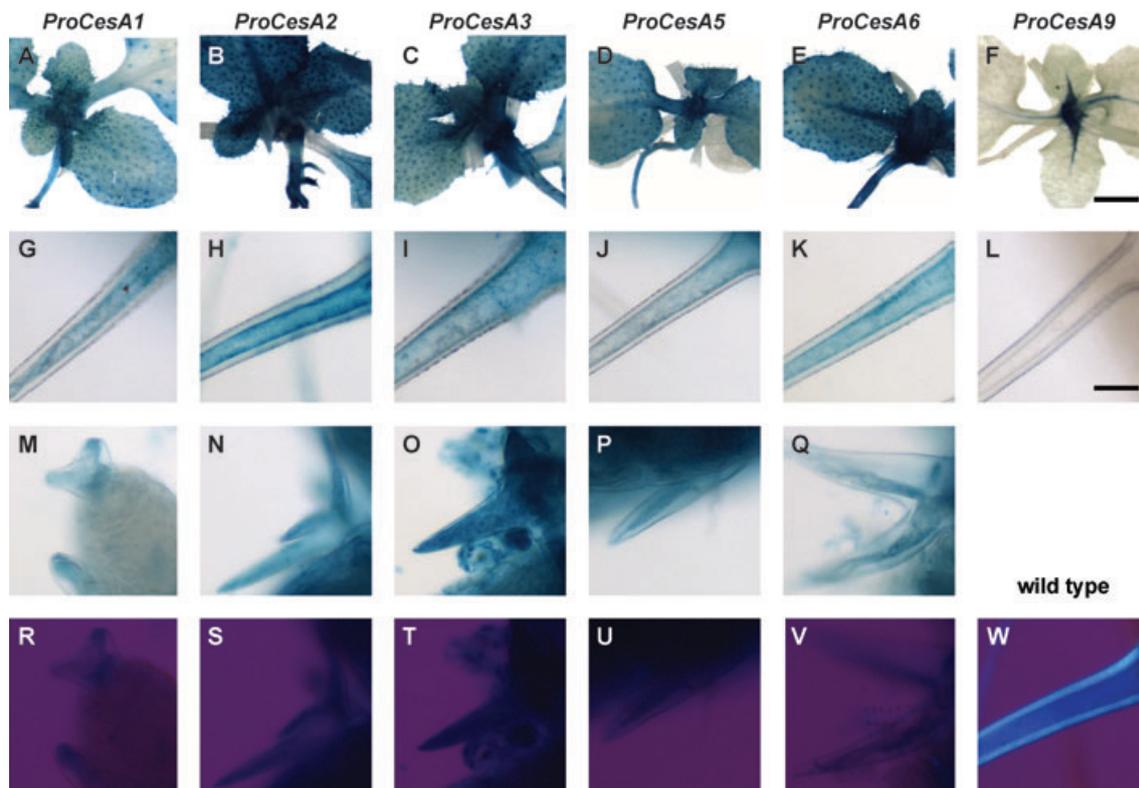
Figure 1. Brightfield images of *GUS* expression patterns (indicated by the blue color) in primary wall-related *ProCesA(1,2,3,5,6,9)::GUS* arabidopsis lines (A–Q) and polarization images of arabidopsis trichomes (R–W).

(A–F) Whole seedlings showing *GUS* expression in expanding root and shoot tissues for most lines, with expression highly restricted to the shoot apex in the *ProCesA9::GUS* line. The times of color development were 1 h (*ProCesA(1,3)*::GUS) or 3 h (*ProCesA(2,5,6,9)::GUS*).
(G–L) Mature, thick-walled trichomes showing GUS activity in all lines except *ProCesA9::GUS*. The thick walls are visible in brightfield microscopy as the colorless bands outlining the optically-sectioned trichomes.
(M–Q) Young, thin-walled trichomes in the early stages of elongation showing GUS activity for the *ProCesA(1,2,3,5,6)::GUS* lines.
(R–V) The same fields in (M–Q) imaged with polarization optics, including a red I retardation plate.
(W) The birefringent walls of a mature wild type trichome appeared bright blue, which was not the case for the young trichomes with only primary cell walls (R–V). Magnification bars: for A–F, a 2 mm bar is found in (F); for G–W, a 25 μm bar is found in (L).

Relative to the low level of *AtCesA7* gene expression in isolated trichomes and processed shoots, only five of the primary wall-related *CesAs, AtCesA(1,2,3,5,6),* were substantially up-regulated in both samples. Little expression of secondary wall-related *CesA* genes, *AtCesA(4,7,8)*, was detected by qPCR (**Figure 3A**). The qPCR results were compared to the results of microarray analysis using the same trichome isolation protocol (Marks et al. 2009); these results were also normalized to the low value for *AtCesA7* in each sample (**Figure 3B**). The microarray data corroborated the qPCR results for all ten *CesAs* assayed (compare **Figure 3A,B**). The low expression of the secondary wall related *CesAs* as detected by qPCR and microarray analysis in processed shoots seemed inconsistent with *GUS* expression in the xylem of Pro*At(CesA4,7,8)::GUS*

lines. This discrepancy is likely explained by the minor amounts of expanding leaves/elongating stems active in xylem differentiation in the processed shoots used for RNA isolation. In contrast, young regions of the shoots could be selectively analyzed in microscopic assays for GUS activity.

Analysis of trichome secondary wall thickening in arabidopsis lines with mutations in secondary wall *CesA* genes

The lack of secondary wall *CesA* expression in isolated trichomes was unexpected in light of the essential roles of *AtCesA(4,7,8)* in secondary wall cellulose synthesis in arabidopsis xylem (Turner and Somerville 1997; Taylor et al.
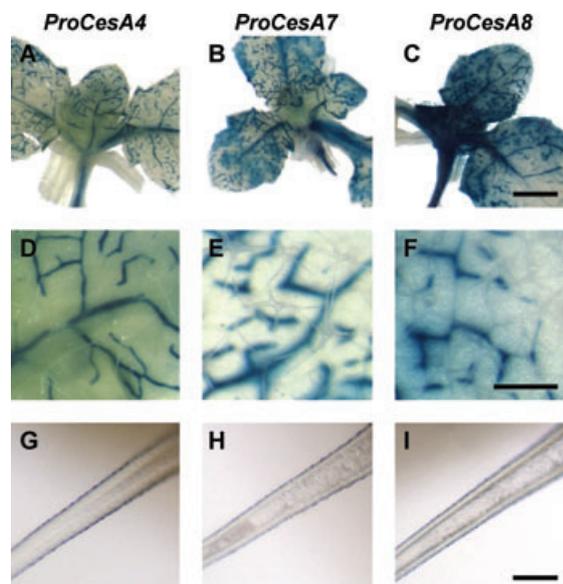
Figure 2. Brightfield images of *GUS* expression patterns (indicated by blue color) in secondary wall-related *ProCesA(4,7,8)::GUS* arabidopsis lines.

(A–C) Whole seedlings after 3 h color development showing tissue-specific GUS expression in shoot and root xylem.
(D–F) Leaves showing GUS expression in xylem.
(G–I) Trichomes showing lack of GUS expression. Magnification bars: for A–C, a 2 mm bar is found in (C); for D–F, a 500 μm bar is found in (F); for G–I, a 25 μm bar is found in (I).

2003; Taylor et al. 2004). To test the role of members of the S1-S3 clades in trichome differentiation, we analyzed three mutant lines of arabidopsis (*irx1-1, irx3-1, irx5-1*) with reduced cellulose content, weak inflorescence stems, and the irregular (or collapsing) xylem phenotype arising from secondary wall cellulose deficiency. In the *irx1-1* mutant the third conserved aspartate residue in *AtCesA8* is replaced by an asparagine residue (D683N), leading to an expected null allele (Taylor et al. 2000). The *irx3-1* mutant has a W859Stop in *AtCesA7*, which causes the protein to be shortened by 168aa (Taylor et al. 2000). The *irx5-1* mutant has a 5kb Ds insertion at L858 in *AtCesA4* (Taylor et al. 2003). The wild type arabidopsis line, Landsberg erecta, was analyzed as a control since these *irx* mutants were in this background. Consistent with the low expression of *AtCesA(4,7,8)* in wild type trichomes, the leaf trichome cell walls in the three *irx* mutants were similar to wild type as shown by the blue color observed under polarization optics with the red 1 retardation plate (**Figure 4A–D**). **Figure 1R–V** illustrate that this blue color does not appear in juvenile trichomes before secondary wall thickening occurs.

Comparison of arabidopsis and cotton gene expression sets

The emerging picture from these experiments was that both branch elongation and secondary wall thickening in shoot trichomes were accomplished by *AtCesA(1,2,3,5,6)*, the members of the P1-P3 clades required for primary wall deposition. This outcome was more comprehensively evaluated by comparing arabidopsis gene identifiers held in common between several gene expression sets (**Table 3**, Supporting Table S2). The six sets of genes used in the comparison were either: (1) strongly co-expressed in arabidopsis with P1-3 *CesAs* (Persson et al. 2005); (2) strongly co-expressed in arabidopsis with S1-3 *CesAs* (compiled from Brown et al. 2005 and Persson et al. 2005); (3) up-regulated ≥ 2-fold in arabidopsis secondary xylem compared to other tissues (Zhao et al. 2005); (4) up-regulated ≥ 2-fold in arabidopsis trichomes compared to processed shoots (Marks et al. 2009); (5) up-regulated ≥ 2-fold in cotton fiber during primary wall deposition (elongation stage at 10 DPA) as compared to secondary wall deposition (thickening stage at 20 DPA); and (6) the converse of set five, which defined a set of genes up-regulated during secondary wall deposition. Since sets 5 and 6 were cotton fiber microarray data, the closest arabidopsis homologs of the up-regulated cotton genes were first identified by BLAST (Supporting Table S1). This resulted in six sets of arabidopsis gene identifiers that could be compared using JMP software (SAS, Cary NC) in order to determine gene identifiers held in common, or intersecting, between them. Percentages of intersection between gene sets (**Table 3**) were based on the 46 or 52 arabidopsis genes, respectively, that were highly co-expressed with primary or secondary wall *CesAs*. For example, 7 genes among those up-regulated in arabidopsis trichomes were held in common with 52 genes co-expressed with secondary wall CesAs (a 13.5% intersection between sets), whereas 22 were held in common with 46 genes co-expressed with primary wall CesAs (a 47.8% intersection between sets). Additional details about the sets of genes compared and the computational and statistical procedures are available in the methods section.

Summarizing the comparisons, for the set of arabidopsis genes co-expressed with primary wall-related (P1-3) *CesA* genes, 47.8% ($P = 2.7E-07$) intersected with the genes up-regulated in isolated trichomes vs. processed shoots. This could be a minimal intersection set since shoots express primary wall-related genes, some of which might not be up-regulated in trichomes. This possibility was investigated further by additional analysis of gene expression in isolated trichomes. The 13,917 genes called present in the isolated trichomes (Marks et al. 2009) included 45 of 46 genes co-expressed with primary wall *CesAs* (a 97.8% intersection) and 19 of 52 genes co-expressed with secondary wall *CesAs* (a 36.5%

Table 1. Gene identifiers and primers used for qPCR in cotton

NCBI accession number	Cotton gene name[a]	Closest arabidopsis homolog	Primers, (F) Forward and (R) Reverse
U58283	GhCESA1	AtCesA8; At4g18780	(F) TGGACTACCCGGTGGATAAGGT
			(R) CTTTCTTGCAAAGTCGGCTGTT
U58284	Gh(AtCesA4-like)	AtCesA4; At5g44030	(F) GCAGCAGACGATACAGAATTCG
			(R) CGTTGTTGATTGCGTCTGAAAC
CO496524	Gh(AtCesA7-like)	AtCesA7; At5g17420	(F) GCCGGCAATCTGTTTACTTACC
			(R) GCTCGAGAATACCAGTTGCAAA
ES812699	Gh(AtCesA1-like)	AtCesA1; At4g32410	(F) CAGTGGATTACCCAGTGGACAA
			(R) GGCACCCACTTTCTTAGCAAAC
AY291285	GhCTL1	AtCTL2; At3g16920	(F) CCGACCAAGAACGACACGTT
			(R) ACCTCGCCCACAAACTTGAT
AF527943	GhCHIA7	AtPOM1/ AtCTL1; At1g05850	(F) GAGCTGCTGGCGATGGTATAA
			(R) GGGTCATCCACCTCCAGATG
CO496020	Gh(AtCOBL4-like)	AtCOBL4; At5g15630	(F) CTTTGATTACAAGCCCCTCGTT
			(R) ATGGCCCAGCTTCCATTAATAG
ES791793	Gh(AtCOB-like)	AtCOB; At5g60920	(F) CGTGAATTGCCCGACATGT
			(R) AAGCAGAGGTGGTGCATTGG
ABV27477	GhFLA6	AtFLA11; At5g03170	(F) CGGTGTGCTTAGCTGGTTGTT
			(R) CACTTCCATCCTCGAAAAAGGT
CO492947	Gh(AteIF-5-like)	putative eIF-5; At1g36730	(F) GGTTGCCATTGTGCAAGGA
			(R) CCGTAGGTGAGCGTTAATCAGA

[a] Gh refers to *Gossypium hirsutum* and At refers to arabidopsis. *GhCESA1, GhCTL1, GhCHIA7,* and *GhFLA6* are established names for full-length cotton genes. Names for cotton ESTs or other partial sequences are derived from their closest match by reciprocal BLAST in the arabidopsis proteome and/or phylogenetic analysis, e.g. *Gh(AtCesA4-like)* (see also the graphs in **Figure 5**).

intersection) (data not shown). This comparison confirmed that primary wall-related processes predominate in trichome differentiation. Surprisingly, only 10.9% of the genes co-expressed with primary wall *CesAs* intersected with genes up-regulated in elongating (10 DPA vs. 20 DPA) cotton fiber. However, the first phase of cotton fiber wall thickening occurs via additional synthesis of primary wall-like components (Meinert and Delmer 1977). Consequently, the expression of many primary wall-related genes had not yet started to decline in cotton fiber at 20 DPA even though genes required for secondary wall deposition were beginning to show increased expression (see **Figure 5A,B**).

The genes co-expressed with primary wall *CesAs* in arabidopsis did not intersect significantly with genes up-regulated for cotton fiber secondary wall deposition (4.4%) or in arabidopsis xylem (2.2%). Conversely, the set of 52 genes co-expressed with secondary wall-related *CesA* genes in arabidopsis had the greatest intersection with genes up-regulated for cotton fiber secondary wall deposition (57.7%) and in arabidopsis xylem (65.4%) ($P = 0$ for both comparisons). The intersection of the secondary wall *CesA* co-expression set with genes up-regulated in arabidopsis trichomes (13.5%) and for cotton fiber elongation (3.9%) was not significant. Supporting Table S2 shows the identity of the genes in the intersection sets.

Expression patterns of primary and secondary wall-related genes in cotton fiber

Toward the end of elongation, the expression of genes required for primary wall synthesis in cotton fiber is down regulated (for examples see Arpat et al. 2004; Hovav et al. 2008; Singh et al. 2009a; Al-Ghazi et al. 2009). This large scale phenomenon was confirmed by qPCR analysis of cotton homologs of *At-COBRA* and *AtCesA1*, *Gh(AtCOB-like)* and *Gh(AtCesA1-like)* respectively, which have been proven by genetic analysis to be involved in primary wall cellulose synthesis in arabidopsis. Both *Gh(AtCOB-like)* and *Gh(AtCesA1-like)* were down-regulated at the end of elongation (**Figure 5A**). *GhCHIA7*, a cotton homolog of *AtCTL1* that has an unknown role in primary wall cellulose synthesis in arabidopsis, was also tested. However, its expression level was relatively high during both primary and secondary wall synthesis in cotton fiber (**Figure 5A**). In addition *GhCTL1*, a secondary wall-related isoform in the same chitinase-like gene family, was up-regulate during secondary wall deposition in cotton fiber (**Figure 5B**; Zhang et al. 2004).

In addition to this secondary wall-related chitinase-like gene, genes encoding secondary wall-related isoforms of CesA and COBRA-like protein were up-regulated for secondary wall deposition in cotton fiber similarly to their expression in xylem

Table 2. Gene identifiers and primers used for qPCR in arabidopsis

Gene name	Locus ID	Primers, (F) Forward and (R) Reverse
AtCesA1	At4g32410	(F) CTCGCTGAAACTGGAGATGTCTT
		(R) CATTGAGGGCAACACTGAGTTC
AtCesA2	At4g39350	(F) CCGGTTTGTAGACCATGCTATG
		(R) TCTCCATCAACCCGTGGACTA
AtCesA3 (a)	At5g05170	(F) GCAAGACTGTTGATGGAGATCGT
		(R) GCACTGAGGACAAGATTGATTCC
AtCesA3 (b)		(F) GGTAGTCCTGCTATTCCTGGTGATA
		(R) CTCTCCCTTCCCACGAGTAAGA
AtCesA4	At5g44030	(F) GCTCAGTGTACCTCGCCAT
		(R) CATCGCCACAGACTTTGCAA
AtCesA5 (a)	At5g09870	(F) GGAGTGGCCTTGAATCTGAAAC
		(R) TTCAACGTCCTCCTCTCCATAAG
AtCesA5 (b)		(F) CGTTACAAGCGCATCAAAGG
		(R) TCAGATTCAAGGCCACTCCTACT
AtCesA6	At5g64740	(F) TGGATCTCTGACAATCGAGTTGAG
		(R) GATTAACCGACCACCGGTGTT
AtCesA7	At5g17420	(F) CCATTGTGGTGATTTGGTCAGT
		(R) TCAGCAGTTGATGCCACACTTGGA
AtCesA8	At4g18780	(F) CGGCAATCCTTACGATGAGAA
		(R) GAATCCCTGAATCCTGAGAAGTGT
AtCesA9	At2g21770	(F) TTCAGCTTCACCAGGCTCTGA
		(R) CTGTTCCCCAAACCCGTAGAT
AtCesA10	At2g25540	(F) AACAAGGCAAGACTCCCACATC
		(R) GCGTAGGAATCTCTCCAGAAACC

(**Figure 5B**). This finding has been confirmed more broadly in microarray analysis of cotton fiber (see Supporting Table S1, Al-Ghazi et al. 2009, and a general discussion in Haigler et al. 2009). Although *GhFLA6* showed earlier up-regulation, expression of most of the secondary wall related genes increased at 19 DPA and reached a sustained high level about 22 DPA, coinciding with cessation of elongation and onset of high rate cellulose synthesis for secondary wall deposition. As shown by phylogenetic analysis (**Figure 6**), these secondary wall related cotton *CesA* genes assayed by qPCR (AAB37767, called *Gh(AtCesA4-like);* CO496524, called *Gh(AtCesA7-like);* and *GhCESA1)* are the closest known *G. hirsutum* homologs of the *CesA* genes defining the S1, S2, and S3 clades in arabidopsis, *AtCesA(4,7,8).* The CesA sequences in the phylogram were edited using uniform criteria described in the methods section, and they include ones from: (a) divergent lineages of plants with fully sequenced genomes (dicot arabiopsis and monocot rice); and (b) cotton, as analyzed previously by phylogenetics (Kim and Triplett 2007) and/or identified by reciprocal BLAST analysis for selection of qPCR targets (Singh et al. 2009a,b; **Figure 5A,B**). Similarly, as determined by reciprocal BLAST analysis, *GhCTL1*, *Gh(AtCOBL4-like),* and *GhFLA6* are the closest known cotton homologs of secondary wall-

related chitinase-like, cobra-like, and arabinogalactan proteins in arabidopsis. The limited knowledge of the mechanistic roles of these proteins in secondary wall cellulose synthesis has been reviewed (Somerville 2006).

Discussion

The classical systems for analyzing cell wall biosynthesis in arabidopsis have been primary walls in the epidermis or cortex and secondary walls in the xylem (vessels and fibers) or cortex (interfascicular fibers). Secondary walls in eudicots are often equated with the wall thickenings of xylem cells and cortical fibers, which contain substantial amounts of highly organized cellulose embedded in xylan, lignin, and other minor wall components. A more general definition of secondary cell walls, however, encompasses any thick wall deposited in (mainly) non-expanding regions of the cell. These thick walls may contain cell wall components of multiple types and organization (Evert 2006). For example, the secondary walls in shoot trichomes and cotton fiber are structurally distinct. The mainly cellulosic commercial cotton fiber is covered by only a thin, cuticulated primary wall (Meinert and Delmer 1977). In contrast, the cell walls of arabidopsis shoot trichomes include abundant pectin, cellulose, mannose-containing polysaccharides, lignin, and cuticle/wax appropriate for their defensive role and the general role of the epidermis in regulating plant water status. Cellulose is not the dominant polymer in arabidopsis shoot trichomes; instead, pectin is > 2-fold more abundant and mannose-containing polysaccharides exist in similar quantity compared to cellulose (Marks et al. 2008). From this perspective, it cannot be assumed that the biosynthetic processes for all secondary walls mimic those of the frequently studied xylem cells and cortical fibers. Similar arguments were made in the introduction about the convergent evolutionary processes giving rise to trichomes, classically defined as all elongated epidermal cells.

The research reported here aimed to determine the extent to which secondary wall deposition in arabidopsis shoot trichomes is an appropriate model for cotton fiber secondary cell wall deposition. Contrary to this possibility, we show here that cotton fiber secondary wall deposition is more similar to the process in arabidopsis xylem. First, only the *CesA* genes in the P1-3 clades, *AtCesA(1,2,3,5,6),* are expressed in arabidopsis shoot trichomes throughout their development, as indicated by GUS activity in *ProCesA::GUS* reporter gene lines. The validity of these results is supported by strong GUS activity in expanding shoot tissues, as expected from prior genetic analysis, as well as by the high expression of each of these genes in trichomes and processed shoots as determined by microarray and/or qPCR analysis. Immature trichomes on young leaves in the shoot apex with only thin primary walls show GUS activity from
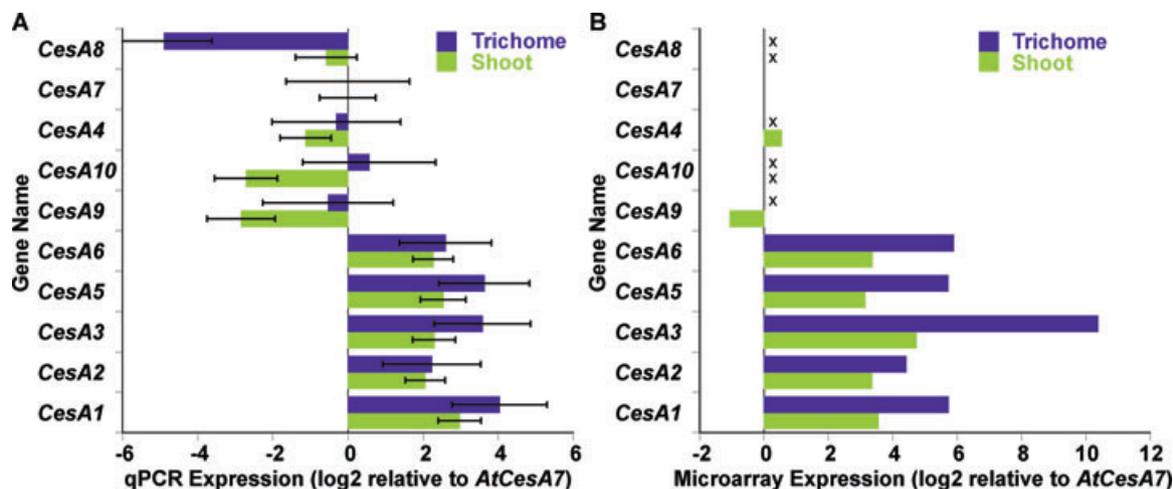
Figure 3. Expression of ten *CesA* genes in isolated arabidopsis trichomes and processed shoots analyzed by qPCR (A) and microarray anaysis (B; Marks et al. 2009).

In each case, log2 data were expressed relative to *AtCesA7*. Both analyses show similar levels of expression of five *CesA* genes in the P1-3 clades, *AtCesA(1,2,3,5,6)*, in isolated trichomes and processed shoots. In contrast, two other P-clade *CesA* genes, *AtCesA(9,10)*, and the entire set of S- clade *CesA* genes, *AtCesA(4,7,8)*, had low (in qPCR) and/or undetectable expression (in the microarray experiment). For qPCR, means for 3-6 replicates are shown along with SE values. For microarray data, an 'x' in place of a histogram bar indicates that the gene was not called present; statistical analysis and error estimations for this experiment are available elsewhere (Marks et al. 2009).

the same set of *ProCesA::GUS* reporter gene lines. Therefore, numerous *CesA* genes in the P1-3 clades, *AtCesA(1,2,3,5,6)*, turn on during the early elongation stage of trichome differentiation. This observation is consistent with data on isolated *sstsim* mutant trichomes with thinner cell walls–although arrested at
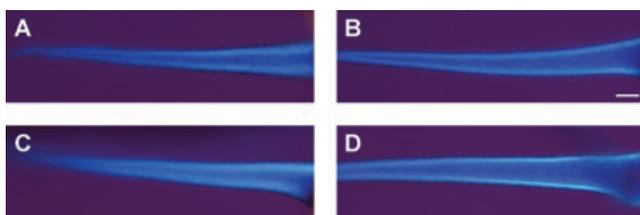

Figure 4. Polarization images showing similar leaf trichome secondary wall deposition in wild type and three arabidopsis lines with mutations in secondary wall related *CesA* genes.

Use of a red I retardation plate caused birefringent secondary walls to appear blue for trichomes of the:
(A) Landsberg erecta wild type;
(B) *irx5-1* mutant in *AtCesA4*;
(C) *irx3-1* mutant in *AtCesA7*;
(D) *irx1-1* mutant in *AtCesA8*.
All trichomes were oriented in the microscope at the same angle that maximized birefringence of the wild type trichomes. Magnification bar, the 25 μm bar in (B) applies to A–D.

an early developmental stage, they express *CesA* genes in the P1-3 clades (Marks et al. 2009). Microarray and qPCR analysis also support the involvement of *CesA* genes in the P1-3 clades in trichome secondary wall thickening. Trichomes were isolated from whole shoots composed mainly of mature leaves and stems prior to RNA isolation for microarray analysis. Therefore, genes expressed during trichome wall thickening should have been well represented. Current evidence supports the continued deposition of cellulose during trichome wall thickening as shown by strong birefringence arising only later in trichome maturation (**Figure 1**) as well as staining consistent with cellulose existing on the innermost layer of thick trichome walls (see Figure 6 in Marks et al. 2008).

Second, *CesA* genes in the S1-3 clades, *AtCesA(4,7,8)*, had low or no expression in trichomes in all assays even though they were expressed as expected in xylem as shown by the *ProCesA::GUS* reporter gene assays. Consistently, mutations in *AtCesA(4,7,8)* did not prevent thickening or development of birefringence in trichome secondary walls even though mutation in any of these genes disrupts secondary wall cellulose synthesis in xylem (Turner and Somerville 1997; Taylor et al. 2003; Taylor et al. 2004). Therefore, *CesA* genes in the P1-3 clades are logical candidates to support cellulose synthesis during the secondary wall thickening of arabidopsis shoot trichomes. Future analysis of the trichomes in arabidopsis lines with mutations in CesAs within the P1-3 clades can test whether particular CesA genes have unique roles in this thickening

Table 3. Percentage intersection between two sets of arabidopsis genes that were co-expressed with either primary or secondary wall *CesA* genes and four other sets of genes with ≥ 2-fold up-regulation in microarray analysis of: arabidopsis trichomes; elongating cotton fiber; thickening cotton fiber; or arabidopsis secondary xylem

Gene expression sets arising from microarray analysis	Primary wall *CesA* co-expression set in arabidopsis (46 genes)	Secondary wall *CesA* co-expression set in arabidopsis (52 genes)
3 539 genes up-regulated ≥ 2-fold in arabidopsis shoot trichomes	47.8%*	13.5%
Arabidopsis homologs of 1 358 genes up-regulated ≥ 2-fold during cotton fiber elongation (10 DPA)	10.9%	3.9%
Arabidopsis homologs of 1 577 genes up-regulated ≥ 2-fold during cotton fiber thickening (20 DPA)	4.4%	57.7%**
852 genes up-regulated ≥ 2-fold in arabidopsis xylem	2.2%	65.4%**

The arabidopsis homologs of cotton genes were determined to allow comparison with the other sets of arabidopsis genes. Percentage calculations were based on the 46 or 52 arabidopsis genes that were highly co-expressed with primary or secondary wall *CesA*s, respectively. For example, 47.8% of the 46 arabidopsis genes co-expressed with primary wall *CesA*s were found in the set of 3 539 genes up-regulated ≥ 2-fold in arabidopsis trichomes. Asterisks indicate the probability of greater intersection between the gene sets than would occur by chance (*, P = 2.7 E-07; **, P = 0)

process. Contrary to the results for trichomes, cotton fibers showed up-regulation of orthologs of *AtCesA(4,7,8)*-*Gh(AtCesA4-like)*, *Gh(AtCesA7-like),* and *GhCESA1*, respectively-at the onset of secondary wall deposition.

Third, bioinformatic comparison of gene expression sets confirmed that gene expression in secondary wall stage cotton fiber was more similar to arabidopsis secondary xylem, not shoot trichomes. Gene sets used in this comparison were two sets of genes previously determined to be strongly co-expressed with *CesA* genes in the P1-3 or the S1-3 clades (Brown et al. 2005; Persson et al. 2005) and microarray data from: (a) arabidopsis isolated trichomes (Marks et al. 2009);
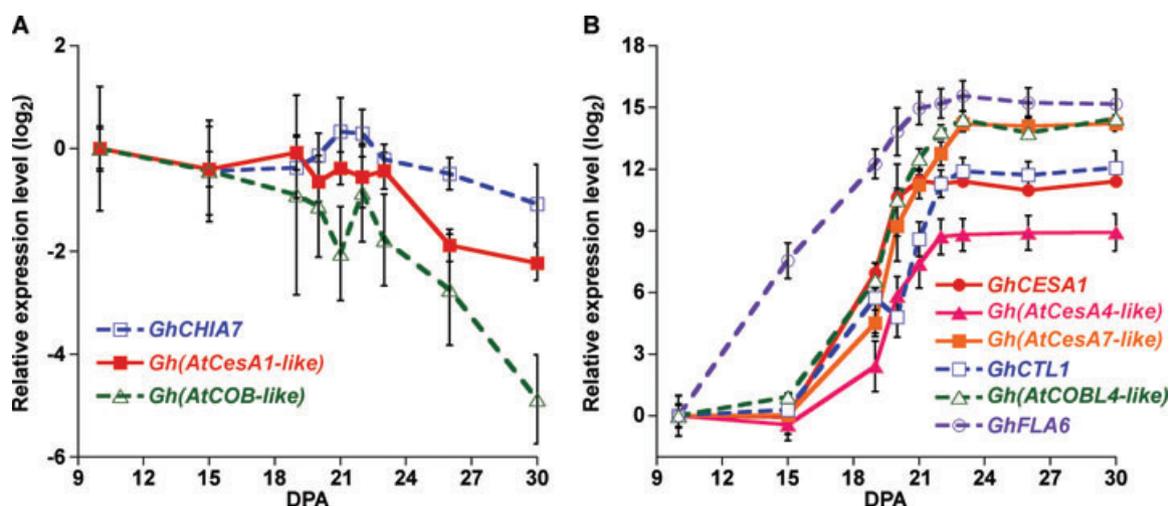

Figure 5. Patterns of expression assayed by qPCR in 10-30 DPA cotton fibers for putative cotton orthologs of primary wall (A) and secondary wall (B) related genes.

The expression profile for each gene is shown relative to its level at 10 DPA, a time of high-rate elongation. The transition between primary and secondary wall deposition occurred between 17-20 DPA. Each data point is the mean ± *SE* value from three biological replicates. The graphs for *GhCTL1*, *GhCESA1*, and *Gh(AtCesA4-like)* are republished from Singh et al. (2009b) with kind permission from Springer Science + Business Media (License Number 2237680337315).
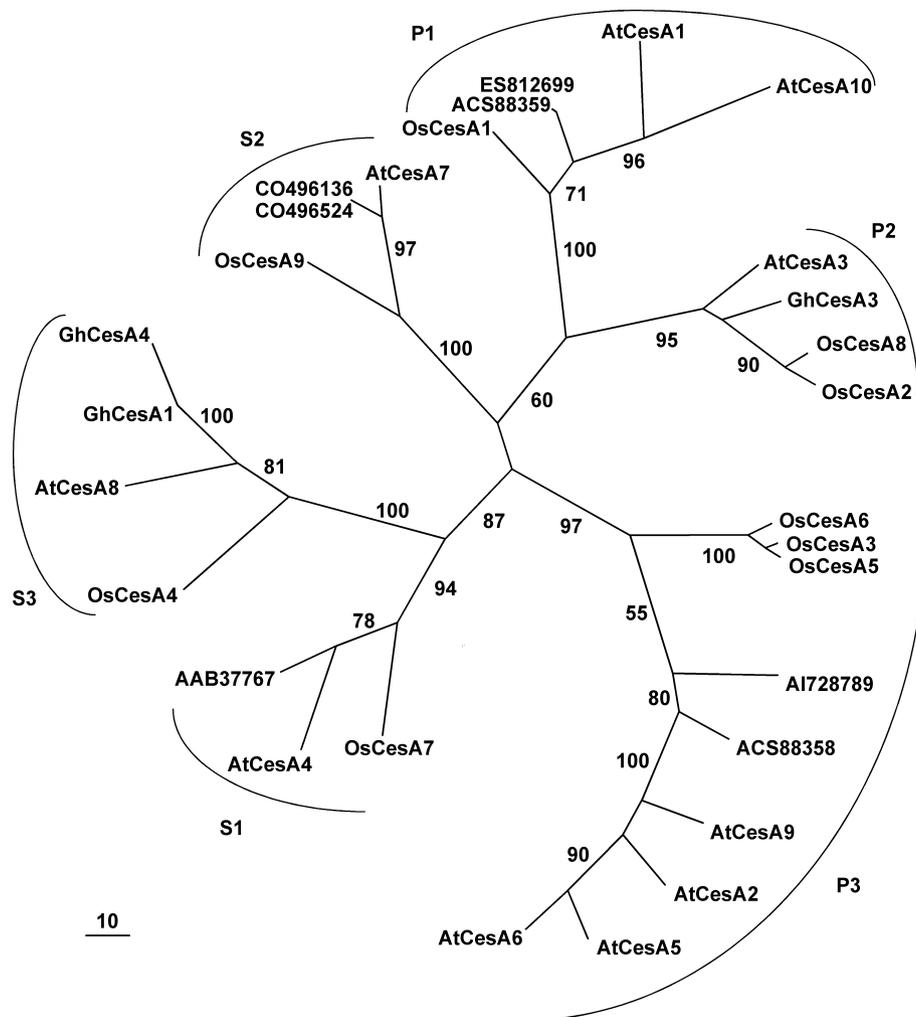
Figure 6. Parsimony phylogram corresponding to the majority consensus of 1000 bootstrap replicates of deduced CesA amino acid sequences from arabidopsis and rice, as well as selected sequences from cotton.

The analysis supports putative orthology between cotton sequences used as qPCR targets and arabidopsis and rice sequences in the primary (P1–3) and secondary (S1–S3) cell wall-related clades of CesAs. All sequences were edited as described in methods. Sequence identifiers are as follows. Arabidopsis, At, TAIR locus I.D.s: At4g32410, At4g39350, At5g05170, At5g44030, At5g09870, At5g64740, At5g17420, At4g18780, At2g21770, At2g25540). Rice, Os, TIGR locus I.D.s: Os05g08370, Os03g59340, Os07g24190, Os01g54620, Os03g62090, Os07g14850, Os10g32980, Os07g10770, Os09g25490. Cotton, Gh, GenBank accession numbers: AAB37766, AAB37767, AAD39534, AF413210, AI728789, ES812699, CO496136, CO496524, ACS88359, and ACS88358.

(b) arabidopsis secondary xylem (Zhao et al. 2005); and (c) cotton fiber during times when it was elongating via primary wall synthesis (10 DPA) or beginning to thicken via secondary wall synthesis (20 DPA) (Supporting Table S1). The new microarray results for cotton fiber were consistent with previous studies on cotton fiber using the same or different chips (Gou et al. 2007; Hovav et al. 2008; Singh et al. 2009a), and they were validated by qPCR of six genes with enhanced expression during fiber secondary wall deposition (**Figure 5B**). Other qPCR results (**Figure 5A**) also corresponded with the microarray data

in showing no change in expression at 20 DPA for three genes related to primary wall synthesis, which is consistent with persistent primary wall synthesis during the earliest stage of cotton fiber thickening (Meinert and Delmer 1977). **Table 3** shows that the transcriptional profiles of both thickening cotton fiber and xylem have the most commonality (57.7% and 65.4%, respectively) with genes that were co-expressed with secondary wall *CesA* genes, whereas isolated shoot trichomes have the most commonality with genes that were co-expressed with primary wall *CesA* genes (47.8%). Examining particular

genes in the intersection sets (Supporting Table S2), shows that *AtCesA1,3,6* (the P1-3 isoforms) are found only among genes up-regulated in arabidopsis trichomes. *AtCesA4,7,8* or their cotton homologs (the S1-3 isoforms) are found in two gene sets—those up-regulated in arabidopsis xylem and secondary wall stage cotton fiber. The greater similarity of cotton fiber secondary wall stage gene expression to xylem extends to many other genes proven by genetic analysis to be required for secondary wall cellulose synthesis in arabidopsis xylem and/or cortical fibers (Supporting Table S1 and S2; Haigler et al. 2009). Interestingly, this similarity is deepened by evidence that at least some cultivars of cotton synthesize 0.37-1.08% lignin-related phenolics in addition to other phenolic molecules during cotton fiber differentiation (Fan et al. 2009).

Cumulatively, the results show that *CesA* genes in either the primary (P1-3) or secondary (S1-3) clades can be deployed flexibly to support secondary wall thickening in different types of trichomes. If primary and secondary walls are defined simply as expanding vs. thickening walls, the concept of primary wall-related as contrasted with secondary wall-related CesAs needs revision. Further work may reveal structural differences in cell walls that are deposited by proteins from the P1-3 clades defined by *AtCesA(1,3,6 or 6-like proteins)* vs. the S1-3 clades defined by *AtCesA(4,7,8)*. This is a worthwhile goal for future research because uncovering potential common features of cellulose synthesized by different types of CesA proteins will aid the understanding of how particular cell wall properties are regulated at the nanoscale (Haigler and Roberts 2009). It is possible that cell wall types would be best defined by the clades of CesA proteins involved in their biosynthesis. The results also indicate that arabidopsis xylem and cortical fibers are preferred models for functional testing of putatively orthologous genes that may regulate cotton fiber secondary wall cellulose synthesis. However, use of model systems to understand specialized cell types in crop plants may have limited value given that different genes/proteins within a "cell wall toolbox" are likely to be used to generate the diverse cell wall physical properties that confer their specialized structures, natural functions, and economic usefulness. In addition, about 30% of cotton genes have no recognizable arabidopsis homologs (with 1E-5 cut-off; Rong and Paterson 2009), implying that research on cotton fiber itself will be required to provide novel insight into the regulation of fiber maturity and strength via synthesis of a secondary wall containing > 90% cellulose.

Materials and Methods

Accession numbers of genes and germplasm lines and primers for qPCR

Table 1 shows the accession numbers and primers for cotton genes assayed by qPCR in fiber of *Gossypium hirsutum* cv.

Deltapine90. The cotton gene names used in this article are as follows: (a) *GhCESA1, GhCTL1, GhCHIA7,* and *GhFLA6* are established names for full-length cotton genes; and (b) ESTs or other partial sequences were named for their closest match by reciprocal BLAST in the arabidopsis (*Arabidopsis thaliana*) proteome, e.g. *Gh(AtCesA4-like)*. For the CesA family emphasized in this report, putative orthology between cotton qPCR targets and arabidopsis family members was confirmed by phylogenetic analysis (**Figure 6**).

For 10 *CesA* genes assayed by qPCR in arabidopsis, the locus identifiers and primers are listed in **Table 2**. Transgenic arabidopsis lines (Columbia background) expressing one of ten *CesApromoter::GUS* reporter gene constructs were made by D. Delmer and coworkers (available with germplasm names CS70755 - 64 as listed in the Arabidopsis Information Resource (TAIR, http://www.arabidopsis.org). These lines are named *ProCesA#::GUS*, where # corresponds to 1-10 of the CesA name. Here, we use parentheses to refer to multiple line numbers exhibiting similar behavior, e.g *ProCesA(1,3,6)::GUS*. The lengths of the 5′ fragments cloned from up-stream of the AUG (start) codon for each native *CesA* gene are: *AtCesA1*, 1118 bp; *AtCesA2*, 2462 bp; *AtCesA3*, 2176 bp; *AtCesA4,* 805 bp; *AtCesA5*, 2619 bp; *AtCesA6*, 2472 bp; *AtCesA7*, 1694 bp; *AtCesA8*, 948 bp; *AtCesA9*, 978 bp; and *AtCesA10*, 2634 bp. Additional details about production of the reporter gene constructs are available at TAIR.

For analysis of effects of mutation in secondary wall *CesA* genes on leaf trichome secondary wall thickening, we analyzed the *irx5-1* mutant in *AtCesA4*, the *irx3-1* mutant in *AtCesA7*, and the *irx1-1* mutant in *AtCesA8*. Homozygous seed stocks were obtained from the Arabidopsis Biological Resource Center (http://abrc.osu.edu/) for *irx3-1* (germplasm/stock CS104) and *irx1-1* (germplasm/stock CS18); seeds of *irx5-1* were donated by Simon Turner. As characterized previously (Turner and Somerville 1997; Taylor et al. 2003; Taylor et al. 2004), each of these genes is required non-redundantly for xylem secondary wall cellulose synthesis. Mutation in any one of them results in cellulose deficiency and collapsing xylem elements. These mutations were identified in the Landsberg erecta background, which was analyzed in parallel as a control.

Plant growth conditions and timing of cotton fiber development

Arabidopsis plants were grown from sterilized, vernalized seeds in potting mix (Metro-Mix 360 with nutrients and wetting agent, SUNGRO Horticulture, Bellevue, WA) in a growth chamber with 16 h photoperiod and 22/20°C day/night temperatures and watered 3 times each week by standing pots in dH$_2$O for 4 h.

Cotton plants were grown in a greenhouse with tight temperature control as previously described (Singh et al. 2009a),

and fiber for RNA extraction was harvested in September. The time course of fiber development under these conditions has been published previously, and the transition between primary and secondary wall deposition occurred at 17-22 DPA, with elongation ending and a high rate of cellulose synthesis beginning at 22 DPA (Singh et al. 2009a).

GUS assay

Seedlings about 15 d old were: (a) fixed in ice-cold acetone (90% v/v; 45 min); (b) rinsed (2 × 20 min) in 25 mM phosphate buffer, pH 7.0, containing 1.25 mM $K_3Fe(CN)_6$, 1.25 mM $K_4Fe(CN)_6$, 0.25% (w/v) Triton X-100, and 0.125 M EDTA; (c) vacuum-infiltrated (15 min) with rinsing solution containing 1.25 mg/ml 5-bromo-4-chloro-3-indolyl-β-D- glucuronide, followed by color development at 37°C for 1, 3, and/or 24 h; (d) washed in water (15 min); and (e) stored in 70% EtOH (4°C) until observation.

Microscopic analyses

Micrographs were captured with a digital camera (Model Q5; QImaging, Surrey, BC, Canada). For the GUS assay, whole plants were photographed using a dissecting microscope with a combination of transmitted and incident light. Trichomes were photographed at higher magnification using brightfield and polarization optics, which allowed birefringence of secondary wall cellulose to be detected (Olympus BH-2 polarizing microscope; Olympus Corp., Center Valley, PA). Juvenile trichomes in the early stages of elongation and before wall thickening commenced were identified after dissection of small leaves at the shoot apex. Brightfield images, to show blue color arising from GUS activity in trichomes (or its absence), were captured under identical optical conditions, and the levels function in Adobe Photoshop was adjusted identically across the set to optimize image contrast for viewing. The polarization images were also captured under identical optical conditions for the set, and the stage was rotated for consistent alignment of the trichome branches with the optical axis producing maximal birefringence of the secondary walls. A first-order red (530 nm) retardation plate was used to increase the intensity and color contrast arising from birefringent secondary walls, if present. No image adjustments were made to the polarization micrographs, allowing the intensity of birefringence to be qualitatively compared between samples.

qPCR to determine gene expression profiles

To assay the expression levels of *CesA* genes in arabidopsis, RNA was extracted from three independent replicates of isolated shoot trichomes and processed shoots generally as previously described (Marks et al. 2008). Modifications included

the use of 100 μm glass beads (BioSpec Products, Bartlesville, OK) and retention of trichomes on 70 μm nylon mesh followed by hand-removal of debris. At the start of RNA isolation, processed shoots were ground in a mortar and pestle under liquid nitrogen. Trichomes (about 1 ml packed cell volume) were disrupted in a microfuge tube by agitation with 3 sterilized 1.5 mm beads (BioSpec Products) and 100 μl Buffer RLT (RNeasy Mini Kit, QIAGEN) using a triturator (Silamat S5; Ivoclar Vivadent AG, Liechenstein). For five cycles, 7 sec in the triturator was followed by 7 sec in liquid nitrogen. DNA was removed (On-Column DNase Digestion, SIGMA-ALDRICH, St. Louis, MO) before qPCR. Specific primers for the 10 arabidopsis *CesA* genes were designed using Primer Express V3 software (Applied Biosystems, Foster City, CA) except that two previously published sequences were used [*AtCesA4* Forward (Taylor et al. 2003) and *AtCesA7* Reverse (Zhong et al. 2003)] (**Table 2**). Primers were designed as before (Singh et al. 2009a) with parameters 45-55 GC%, 18-25 nucleotide length, 58-60°C Tm, and 100-150 bp amplicon length. Primers were checked by BLAST against all arabidopsis genes at TAIR to ensure that they targeted one *CesA* gene. The qPCR assay was performed as described (Singh et al. 2009a), except that 200 ng RNA was used to make cDNA and 45 cycles were used instead of 40. Two sets of primers were used to amplify *AtCesA3* and *AtCesA5* with similar results, and the results are averaged. Amplification of At3g1260, a DNA-binding protein with similar expression in trichomes and shoots, was used as an endogenous control as described previously (Marks et al. 2008). Each value was expressed relative to the level of *AtCesA7* in the same sample because this gene had a low expression level in both processed shoots and trichomes as determined by qPCR analysis reported here as well as in microarray experiments (Marks et al. 2009).

For comparing secondary wall deposition-related genes in arabidopsis and cotton, the best matches to putatively orthologous arabidopsis sequences were identified in ESTs from the *Gossypium* genus available at the National Center for Biotechnology Information (NCBI). BLAST analysis was used to show the reciprocal best match of the translated protein sequences (including any longer contig that could be identified at http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/gimain.pl?gudb=cotton) in both species. Primers were designed as described above and checked by BLAST against all cotton sequences at NCBI for amplification of only one known cotton sequence. Because there is no complete cotton genome sequence, it is possible that multiple closely related genes were amplified. Genes assayed and the primer sequences used for this experiment are shown in **Table 1**. As described previously (Singh et al. 2009a), qPCR was performed starting with RNA extracted from isolated fibers of known age. The qPCR results for each gene were: (a) normalized to an endogenous control, *Gh(AteIF5-like)*, a cotton homolog

of elongation initiation factor 5 that was previously validated as a control for cotton fiber qPCR (Haigler et al. 2009); and (b) expressed relative to the value at 10 DPA, a time of high rate of elongation.

Phylogenetic analysis

Cotton peptide sequences included in the phylogram were either full length cotton CesAs in NCBI, identified in NCBI by reciprocal BLAST against the 10 arabidopsis CesAs for selection of qPCR targets assayed in this work, or other partial sequences used previously in phylogenetic analysis of cotton CesAs (Kim and Triplett 2007). Sequences that were very short (<180 aa), contained ambiguous base calls, or were identical to longer sequences were not analyzed. Full length *CesA* sequences from the fully sequenced genomes of rice and arabidopsis were included, and these embrace the two major groups of flowering plants, the Monocotyledonae and Dicotyledonae, respectively. All sequences were edited to remove the N-terminus upstream of the DCDH consensus sequence. Sequences were aligned using Prankster (Löytynoja and Goldman 2008) with gap opening/extension penalties of 1.0000/1.00 and 2 iterations. BioEdit (Hall 1999) was used to remove large gaps and segments of uncertain homology from the alignment (Baldauf 2003). Phylograms were constructed from the aligned sequences using the heuristic search method in PAUP* (version 4.1b10, Sinauer Associates, Sunderland, MA) with all characters given equal weight. The topology was tested with 1 000 bootstrap replicates using the parsimony method.

Comparison of gene expression sets

The arabidopsis gene expression sets providing the basis of the analysis were: (a) 46 genes strongly co-expressed with primary wall *CesA* genes, *AtCesA(1,3,6)* (Persson et al. 2005); and (b) 52 genes strongly co-expressed with secondary wall *CesA* genes, *AtCesA(4,7,8)* (compiled from Brown et al. 2005 and Persson et al. 2005). Four genes that were co-expressed with both primary and secondary *CesA* genes were not included in our comparisons (At5g49720, *KOR1*; At1g80350, *KTN1*; At3g08550, *KOB1*; and At1g65580, *FRA3*).

Using JMP software (SAS, Cary NC), these two sets of genes were evaluated for percentage intersection with four sets of genes with significantly different, ≥ 2-fold, up-regulation in microarray experiments: (a) 3539 genes up-regulated in arabidopsis trichomes vs. processed shoots (Marks et al. 2009); (b) 1 358 genes up-regulated during cotton fiber elongation (at 10 vs. 20 DPA); (c) 1 577 genes up-regulated at the onset of cotton fiber secondary wall deposition (at 20 vs. 10 DPA); and (d) 852 genes up-regulated in arabidopsis secondary xylem (containing vessels and fibers) vs. non-vascular tissue (Zhao et al. 2005).

As referenced above, the microarray analysis of isolated arabidopsis trichomes and secondary xylem tissue have been published. The cotton fiber microarray analysis used 10 and 20 DPA fiber from greenhouse grown *G. hirsutum* cv. Texas Marker 1, a genetic and cytogenetic standard for domesticated elite cotton. For three biological replicates at each DPA, fiber was separated from developing seeds by agitation with glass beads under liquid nitrogen (adapting methods from Taliercio and Boykin 2007), and RNA was isolated (Wilkins and Smart 1996), treated with DNAase, and linearly amplified (TargetAmp™ 1-Round aRNA Amplification kit, Epicentre Biotechnologies, Madison, WI). A custom high-density chip with 283 000 features (an average of seven distinct, approx. 60 bp probes for each of 42 429 unigenes derived from several cotton species) was used for contracted microarray analysis (Nimblegen Systems, Inc.; Madison, WI). The chip, which is described further at www.cottonevolution.info, was designed to independently interrogate paralogs and included probes for cotton unigenes homologous (with 1E-10 cut-off) to 16 993 unique loci in arabidopsis (J. Udall, pers. comm.). Raw expression values for each unigene represented on the chip were obtained by median polishing the seven redundant probes using Tukey's Biweight estimator (Tukey 1977; Velleman and Hoaglin 1981). In R (R Development Core Team 2005), polished values were natural log transformed, median centered and scale normalized, SAS was used to assess differential expression (Cary, NC), and R was used to control the false discovery rate. Differential expression was indicated by a *q*-value ≤ 0.05. An EXCEL file showing the microarray data for 7605 cotton genes with significant differences of expression in fiber between 10 and 20 DPA is provided as Supporting Table S1.

Also in R, Fisher's Exact Probability test was used to determine whether the number of genes in an intersection set was significantly greater than would have occurred by chance (P ≤ 0.05) (Ewens and Grant 2005). The underlying number of arabidopsis loci corresponding to each microarray was used in the test: for the arabidopsis arrays, this was the number of loci represented on the arabidopsis chip; for the cotton arrays, this was the 16 993 arabidopsis loci matched at 1E-10 or better by the cotton unigenes represented on the cotton chip.

References

Al-Ghazi Y, Bourot S, Arioli T, Dennis ES, Llewellyn DJ (2009) Transcript profiling during fiber development identifies pathways in secondary metabolism and cell wall structure that may contribute to cotton fiber quality. *Plant Cell Physiol.* **50**, 1364–1381.

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

Baldauf SL (2003) Phylogeny for the faint of heart: a tutorial. *Trends Genet.* **19**, 345–351.

Boopathi MN, Ravikesavan R (2009) Emerging trends in enhancement of cotton fiber productivity and quality using functional genomics tools. *Biotech. Mol. Biol. Rev.* **4**, 11–28.

Brown DM, Zeef LAH, Ellis J, Goodacre R, Turner SR (2005) Identification of novel genes in *Arabidopsis* involved in secondary cell wall formation using expression profiling and reverse genetics. *Plant Cell* **17**, 2281–2295.

Desai A, Chee PW, May OL, Paterson AH (2008) Correspondence of trichome mutations in diploid and tetraploid cotton. *J. Heredity* **99**, 182–186.

Desprez T, Juraniec M, Crowell EF, Jouy H, Pochylova Z, Parcy F, Höfte H, Gonneau M, Vernhettes S (2007) Organization of cellulose synthase complexes involved in primary cell wall synthesis in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* **104**, 15572–15577.

Dillehay TD, Rossen J, Andres TC, Williams DE (2007) Preceramic adoption of peanut, squash, and cotton in northern Peru. *Science* **316**, 1890–1893.

Djerbi S, Lindskog M, Arvestad L, Sterky F, Teeri TT (2005) The genome sequence of black cottonwood (*Populus trichocarpa*) reveals 18 conserved cellulose synthase (CesA) genes. *Planta* **221**, 739–746.

Doblin MS, Kurek I, Jacob-Wilk D, Delmer DP (2002) Cellulose biosynthesis in plants: from genes to rosettes. *Plant Cell Physiol.* **43**, 1407–1420.

Evert RF (2006) *Esau's Plant Anatomy*, *3rd Ed*, Wiley Interscience, Hoboken, NJ.

Ewens WJ, Grant GR (2005) *Statistical Methods in Bioinformatics: An Introduction*, *2nd ed*, Springer, New York, NY.

Fan L, Shi WJ, Hu WR, Hao XY, Wang DM, Yuan H, Yan HY (2009) Molecular and biochemical evidence for phenylpropanoid synthesis and presence of wall-linked phenolics in cotton fiber. *J. Int. Plant Biol.* **51**, 626–637.

Gou JY, Wang LJ, Chen SP, Hu WL, Chen XY (2007) Gene expression and metabolite profiles of cotton fiber during cell elongation and secondary wall synthesis. *Cell Res.* **17**, 422–434.

Guan XY, Yu N, Shangguan XX, Wang S, Lu S, Wang LJ, Chen LJ (2007) *Arabidopsis* trichome research sheds light on cotton fiber development. *Chin. Sci. Bull.* **52**, 1734–1741.

Haigler CH, Zhang D, Wilkerson CG (2005) Biotechnological improvement of cotton fibre maturity. *Physiol. Plant.* **124**, 285–294.

Haigler CH, Singh B, Wang G, Zhang D (2009) Genomics of cotton fiber secondary wall deposition and cellulose biogenesis. In: Paterson AH, ed. *Genetics and Genomics of Cotton, Plant Genetics and Genomics: Crops and Models 3*, Springer Science +Business Media, LLC, New York. pp. 385–417.

Haigler CH, Roberts AW (2009) Biogenesis of cellulose nanofibrils by a biological nanomachine, In: Lucian L**,** Rojas O, eds. *The Nanoscience and Technology of Renewable Biomaterials*, Wiley-Blackwell Publishing, Oxford UK. 368 pp.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids Symp. Ser.* **41**, 95–98.

Holland N, Holland D, Helentjaris T, Dhugga KS, Xoconostle-Cazares B, Delmer DP (2000) A comparative analysis of the plant cellulose synthase (CesA) gene family. *Plant Physiol.* **123**, 1313–1323.

Hovav R, Udall JA, Hovav E, Rapp R, Flagel L, Wendel JF (2008) A majority of cotton genes are expressed in a single-celled fiber. *Planta* **227**, 319–329.

Jakoby MJ, Falkenhan D, Mader MT, Brininstool G, Wischnitzki E, Platz N, Hudson A, Hülskamp M, Larkin J, Schnittger A (2008) Transcriptional profiling of mature *Arabidopsis* trichomes reveals that NOECK encodes the MIXTA-Like Transcriptional regulator of MYB106. *Plant Physiol.* **148**, 1583–1602.

Kim HJ, Triplett BA (2007) Cellulose synthase catalytic subunit (*Cesa*) genes associated with primary or secondary wall biosynthesis in developing cotton fibers (*Gossypium hirsutum*). In: *Proceedings of the World Cotton Research Conference-4*, International Cotton Advisory Committee, Washington DC, USA. On line at: http://wcrc.confex.com/wcrc/2007/techprogram/P1764.HTM

Ko JH, Beers EP, Han KH (2006) Global comparative transcriptome analysis identifies gene network regulating secondary xylem development in Arabidopsis *thaliana*. *Mol. Gen. Genomics* **276**, 517–531.

Kumar M, Thammannagowda S, Bulone V, Chiang V, Han KH, Joshi CP, Mansfield SD, Mellerowicz E, Sundberg B, Teeri T, Ellis BE (2009) An update on the nomenclature for the cellulose synthase genes in *Populus*. *Trends Plant Sci.* **14**, 248–254.

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

Liu D, Zhang H (2008) Gene cloning: exploring cotton functional genomics and genetic improvement. *Front. Agric. China* **2**, 1–9.

Löytynoja A, Goldman N (2008) Phylogeny-aware gap placement prevents errors in sequence alignment and evolutionary analysis. *Science* **320**, 1632–1635.

Lyon AG, Edwards D (1991) The first zosterophyll from the Lower Devonian Rhynie Chert, Abderdeenshire. *Trans. Royal Soc. Edinburgh.* **82**, 323–332.

Marks MD, Wenger JP, Gilding E, Jilk R, Dixon RA (2009) Transcriptome analysis of *Arabidopsis* wild-type and gl3-sst trichomes identifies four additional genes required for trichome development. *Mol. Plant* **2**, 803–822.

Marks MD, Betancur L, Gilding E, Chen F, Bauer S, Wenger JP, Dixon RA, Haigler CH (2008) A new method for isolating large quantities of *Arabidopsis* trichomes for transcriptome, cell wall and other types of analyses. *Plant J.* **56**, 483–492.

Meinert MC, Delmer DP (1977) Changes in biochemical composition of the cell wall of the cotton fiber during development. *Plant Physiol.* **59**, 1088–1097.

Persson S, Wei H, Milne J, Page GP, Somerville CR (2005) Identification of genes required for cellulose synthesis by regression analysis of public microarray data sets. *Proc. Natl. Acad. Sci. USA* **102**, 8633–8638.

Persson S, Paredez A, Carroll A, Palsdottir H, Doblin M, Poindexter P, Khitrov N, Auer M, Somerville CR (2007) Genetic evidence for three unique components in primary wall cell-wall cellulose synthase complexes in *Arabidopsis. Proc. Natl. Acad. Sci. USA* **104**, 15566–15571.

R Development Core Team (2005) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, http://www.R-project.org

Rong J, Paterson AH (2009) Comparative genomics of cotton and *Arabidopsis*, In: Patterson AH, ed. *Genetics and Genomics of Cotton*, *Plant Genetics and Genomics: Crops and Models 3*, Springer Science + Business Media, LLC, New York. pp. 431–449.

Ryser U (1985) Cell wall biosynthesis in differentiating cotton fibres. *Eur. J. Cell Biol.* **39**, 236–256.

Scheible WR, Eshed R, Richmond T, Delmer D, Somerville C (2001) Modifications of cellulose synthase confer resistance to isoxaben and thiazolidinone herbicides in arabidopsis Ixr1 mutant. *Proc. Natl. Acad. Sci. USA* **98**, 10079–10084.

Serna L, Martin C (2006) Trichomes: different regulatory networks lead to convergent structures. *Trends Plant Sci.* **11**, 274–280.

Singh B, Avci U, Eichler Inwood SE, Grimson MJ, Landgraf J, Mohnen D, Sørensen I, Wilkerson CG, Willats WG, Haigler CH (2009a) A specialized outer layer of the primary cell wall joins elongating cotton fibers into tissue-like bundles. *Plant Physiol.* **150**, 684–699.

Singh B, Cheek HD, Haigler CH (2009b) A synthetic auxin (NAA) suppresses secondary wall cellulose synthesis and enhances elongation in cultured cotton fiber. *Plant Cell Rep.* **28**, 1023–1032.

Somerville C (2006) Cellulose synthesis in higher plants. *Ann. Rev. Cell Dev. Biol.* **22**, 53–78.

Taliercio EW, Boykin D (2007) Analysis of gene expression in cotton fiber initials. *BMC Plant Biol.* **7**, 22–35.

Tanaka K, Murata K, Yamazaki M, Onosato K, Miyao A, Hirochika H (2003) Three distinct rice cellulose synthase catalytic subunit genes required for cellulose synthesis in the secondary wall. *Plant Physiol.* **133**, 73–83.

Taylor NG, Laurie S, Turner SR (2000) Multiple cellulose synthase catalytic subunits are required for cellulose synthesis in arabidopsis. *Plant Cell* **12**, 2529–2939.

Taylor NG, Howells RM, Huttly AK, Vickers K, Turner SR (2003) Interactions among three distinct CesA proteins essential for cellulose synthesis. *Proc. Nat. Acad. Sci. USA* **100**, 1450–1455.

Taylor NG, Gardiner JC, Whiteman R, Turner SR (2004) Cellulose synthesis in the Arabidopsis secondary cell wall. *Cellulose* **11**, 329–338.

Tukey JW (1977) *Exploratory Data Analysis*, 1st Ed. Addison-Wesley, Reading, MA.

Turner SR, Somerville CR (1997) Collapsed xylem phenotype of Arabidopsis identifies mutants deficient in cellulose deposition in the secondary cell wall. *Plant Cell* **9**, 689–701.

Velleman PF, Hoaglin DC (1981) *Applications, Basics, and Computing of Exploratory Data Analysis*. Duxbury Press, Boston, MA.

Wendel JF, Brubaker C, Alvarez I, Cronn R, McD Stewart J (2009) Evolution and natural history of the cotton genus. In: Paterson AH, ed. *Genetics and Genomics of Cotton, Plant Genetics and Genomics: Crops and Models 3*, Springer Science +Business Media, LLC, New York. pp. 3–22.

Wilkins TA, Smart LB (1996) Isolation of RNA from plant tissue. In: Krieg PA, ed. *A Laboratory Guide to RNA: Isolation, Analysis, and Synthesis*. Wiley-Liss, New York. pp. 21–41.

Zhang D, Hrmova M, Wan CH, Wu C, Balzen J, Cai W, Wang J, Densmore LD, Fincher GB, Zhang H, Haigler CH (2004) Members of a new group of chitinase-like genes are expressed preferentially in cotton cells with secondary walls. *Plant Mol. Biol.* **54**, 353–372.

Zhao C, Craig JC, Petzold HE, Dickerman AW, Beers EP (2005) The xylem and phloem transcriptomes from secondary tissues of the *Arabidopsis* root-hypocotyl. *Plant Physiol.* **138**, 803–818.

Zhong R, Morrison WH III**, Freshour GD, Hahn MG, Ye ZH** (2003) Expression of a mutant form of cellulose synthase AtCesA7 causes dominant negative effect on cellulose biosynthesis. *Plant Physiol.* **132**, 786–795.

(Co-Editor: Anne-Mie C. Emons)

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1. EXCEL file of microarray data for all cotton genes (and their arabidopsis homologs) that were significantly

differentially expressed between 10 and 20 DPA in fiber of *G. hirsutum* cv. TM-1. In this file, there is an estimate of difference (log2 ratios) between the two time points that was generated by taking into account the composite variance of the samples, and the fold-change column shows linear values based on those ratios. The expression levels are relative to the value at 10 DPA, so genes with higher expression at 10 or 20 DPA have positive or negative values, respectively.

Table S2. Intersection between gene sets that were co-expressed with primary or secondary wall *CesA* genes and genes that were $\geq$ 2-fold up-regulated in arabidopsis trichomes, elongating cotton fiber, secondary wall stage cotton fiber, or arabidopsis xylem.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.