Plant Biotechnology Journal

Plant Biotechnology Journal (2012) **10**, pp. 301–312 doi: 10.1111/j.1467-7652.2011.00662.x

Overexpression of *GhSusA1* increases plant biomass and improves cotton fiber yield and quality

Yanjie Jiang[1], Wangzhen Guo[1], Huayu Zhu[1], Yong-Ling Ruan[2,3] and Tianzhen Zhang[1,*]

[1]*National Key Laboratory of Crop Genetics and Germplasm Enhancement, Cotton Research Institute, Nanjing Agricultural University, Nanjing, China*
[2]*School of Environmental and Life Sciences, the University of Newcastle, Callaghan, NSW, Australia*
[3]*Australia-China Research Centre for Crop Improvement, the University of Newcastle, Callaghan, NSW, Australia*

Received 14 August 2011;
revised 18 September 2011;
accepted 21 September 2011.
**Correspondence* (fax +86 25 84395307;
email cotton@njau.edu.cn)

Keywords: biomass, cotton, fiber quality, overexpression, sucrose synthase, transgene.

Summary

Cotton (*Gossypium* spp.) is an important economic crop and the largest source of textile fiber in the world. However, to date, only a few genes have been identified that exhibit critical roles in fiber development, and few has shown positive effects on fiber yield and quality in transgenic cotton. Here, we report the characterization of a novel sucrose synthase (*SusA1*) gene from a superior quality fiber germplasm line 7235 in *Gossypium hirsutum*. By association analysis, *GhSusA1* was highly correlated with fiber qualities in (7235× TM-1) recombinant inbred lines based on polymorphism of *GhSusA1* between 7235 and TM-1. Subsequently, based on an interspecific population of 141 BC_1 individuals generated from the cross between TM-1 and *Gossypium barbadense* line, Hai7124, we further mapped GhSusA1 genes on homeologous chromosomes A8 (chro.8) and D8 (chro.24). Suppression of *GhSusA1* in transgenic cotton reduced fiber quality and decreased the boll size and seed weight. Importantly, overexpression of this gene increased fiber length and strength, with the latter indicated by the enhanced thickening of cell wall during secondary wall formation stage. Moreover, increasing *GhSusA1* transcript abundance in vegetative tissues led to elevated seedling biomass. Together, these findings identified *GhSusA1* as a key regulator of sink strength in cotton, which is tightly associated with productivity, and hence a promising candidate gene that can be developed to increase cotton fiber yield and quality.

Introduction

Cotton is the most important textile crop worldwide because of the cellulose-enriched mature fibers. Cotton fibers are single-celled trichomes derived from epidermal cells of the ovule. Sucrose import and metabolism is considered to be a major factor determining sink strength in tissues such as developing cotton seed and fibers (Pugh *et al.*, 2010). In this context, sucrose synthase (Sus, EC 2.4.1.13) is one of only two enzymes which can decompose sucrose into hexoses, by catalysing a reversible reaction but preferentially converting sucrose into fructose and UDP-glucose *in planta* (Chourey and Nelson, 1979; Geigenberger and Stitt, 1993). Its product, UDP-glucose, is known as the most direct substrate for cellulose synthesis. Therefore, by forming a putative complex with cellulose synthase on plasma membrane, Sus has the potential capacity to channel carbon directly from sucrose to cellulose (Amor *et al.*, 1995). Indeed, by using transgenic approach, Ruan *et al.* (2003) demonstrated the critical roles that Sus plays in cotton fiber and seed development (Ruan *et al.*, 2003).

Despite the research progress over the last two decades or so on the role of Sus in cotton fiber and seed development (e.g. Amor *et al.*, 1995; Ruan *et al.*, 2003), the potential impact of overexpression of Sus on the development of cotton fiber, seed or other tissues has not been reported thus far. Therefore, we attempted to fill this knowledge gap by conducting comprehensive molecular genetic and transgenic analyses of Sus expression in cotton.

In this study, we cloned a new *Sus* gene, designated as *GhSusA1,* from a super quality fiber *Gossypium hirsutum* line and found that it was highly correlated with fiber qualities in recombinant inbred lines (RILs) derived from the cross between 7235 and TM-1 lines. By overexpression or suppression of *GhSusA1* in transgenic cotton, we observed positive and negative effects, respectively, on such traits as fiber strength, boll size and seed weight and yield. Our findings demonstrate that *GhSusA1* plays major roles in cotton fiber development and is a key target for improving cotton yield and quality in industrial applications.

Results

Molecular cloning and characterization of the *GhSusA1* gene

The germplasm line 7235 is a high-fiber-quality Upland cotton (Qian *et al.*, 1992). A cDNA library was constructed using fibers from this line at 5–25 days post-anthesis (DPA), covering the period of fiber elongation and secondary cell wall cellulose synthesis. Random sequencing generated 1436 uni-ESTs that were used to construct a microarray to identify differentially expressed ESTs (He *et al.*, 2008). The microarray hybridization identified one up-regulated EST as Sus expressed in fiber mRNAs of 7235 as compared to that from TM-1, the Upland cotton genetic standard (Kohel *et al.*, 1970). Semi- and corresponding real-time quantitative RT-PCR analyses confirmed that

the Sus gene was up-regulated in 7235 compared with that of TM-1 from the fiber elongation at eight DPA to primary–secondary cell wall transition stage of 19 DPA (Figure 1a,b). This time period includes the most important stages of fiber quality formation. The results suggest a positive association between the *GhSusA1* transcript level and fiber quality. Therefore, we cloned the complete 5582-bp genome sequence (GenBank accession number HQ702186) by TAIL-PCR and the 2430-bp ORF cDNA of *GhSusA1* (GenBank accession number HQ702185) from the 7235 line. *GhSusA1* has 15 exons and 14 introns (Figure 1c); while *GhSus1* (U73588) (Ruan *et al.*, 2003) contains 12 exons and 11 introns, and *GhSus2* (FB742813) (Ruan *et al.*, Patent: WO 2008012058-A 31-JAN-2008) has 10 exons and nine introns. The exons of *GhSusA1* are divided at distinct positions, and the intron structure of *GhSusA1* is the most complicated of the known cotton *Sus* genes. This analysis shows that *GhSusA1* is a new *Sus* gene, different from the previously characterized *GhSus1* and *GhSus2*.

At the protein level, sequence analysis showed that GhSusA1 exhibited high level of sequence divergence at the N-terminus as compared with two previously reported *Sus* genes, *GhSus1* (U73588) and *GhSus2* (FB742813) (Figure S1). The *GhSusA1* gene has a 2430-bp ORF encoding a protein of 809 amino acids, which shares a high level of homology with citrus Sus (*CitSUSA*) (Komatsu *et al.*, 2002) with 88% amino acid identity, compared to only 71% with GhSus1 (Figure S1). *GhSusA1* from line 7235 contains 41 putative phosphorylation sites, two putative glycosylation sites targeted by the glycosyltransferase-1 family and a typical 'DxD' consensus region (Figure S1 marked with black triangles) (Christine and Sean, 1998). There are four domains that are highly conserved among sucrose synthesis. In GhSusA1, they are located at 303, 578–580, 655 and 681 amino acid (Figure S1 boxed). By comparing different species and different types of cotton Sus, we found evident differences in the first 80 amino acids although two potential phosphorylation sites are conserved in this region (Figure S1 marked with 'P').

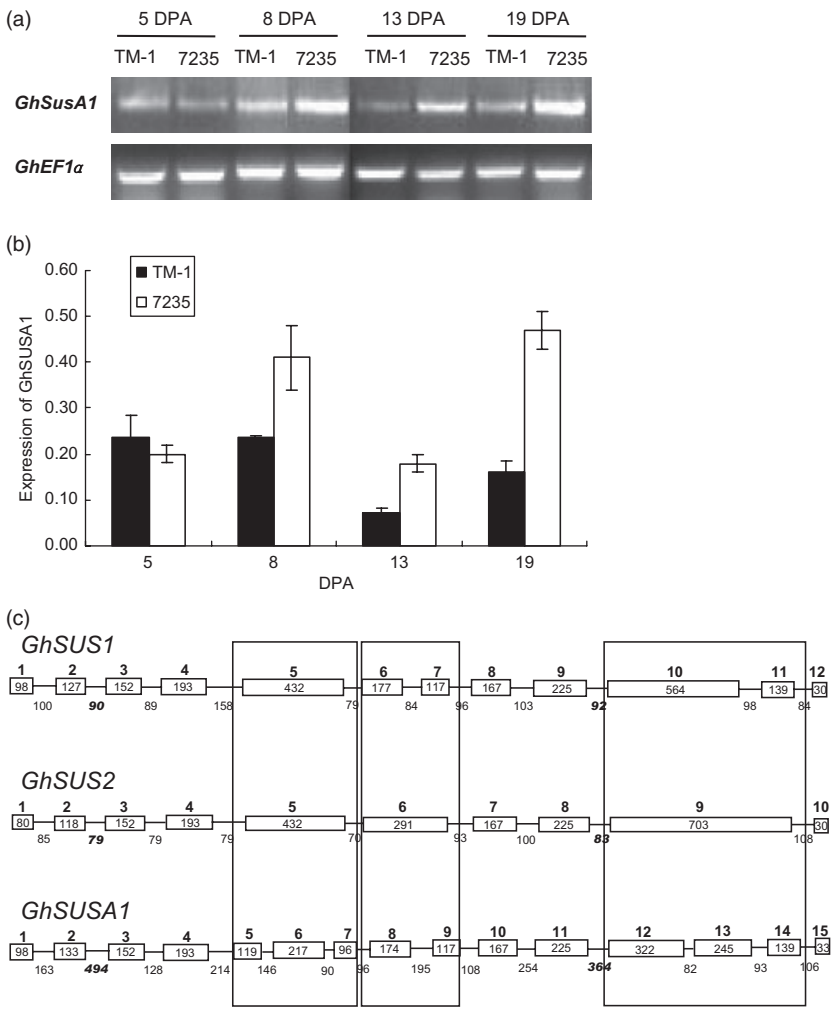

Figure 1 *GhSusA1* is predicted to encode a novel sucrose synthase protein. (a) Expression analysis of *GhSusA1* in fibers of 7235 and TM-1 during fiber elongation (5, 8 and 13 days post-anthesis, DPA) and early secondary wall formation stage (19 DPA). Transcript level of *GhSusA1* was high in the elite quality cotton 7235 than in standard line from 8 DPA afterwards by semi-quantitative RT-PCR. (b) Real-time quantitative RT-PCR of corresponding samples in (a). Vertical bars represented standard deviation (SD). (c) Exon/intron structures of three sucrose synthase genes *GhSus1*, *GhSus2* and *GhSusA1* from cotton. The main differences in exon structure and intron length were boxed and marked bold, respectively. The exons of *GhSusA1* were split at distinct positions, and the introns structure of *GhSusA1* was the longest and most complicated.

A phylogenetic dendrogram of plant Sus showed that GhSusA1 is completely different from the GhSus1 and GhSus2 (Figure S2). They share 70% or less identity in protein level (Figures S1 and S2). *GhSusA1* is obviously different from the *GhSus1* and *2* in encoded amino acid sequence and genomic structure, which probably confers to its different functions or regulations in cotton.

To determine the biochemical functionality of GhSusA1, we expressed it in *Escherichia coli* BL21. The purified protein from bacterial lysates possessed enzyme activity of catalysing sucrose to UDP-glucose and fructose, validating that it functions as a sucrose synthase and has *in vitro* activity (Figure S3).

Association analysis and genome localization of *GhSusA1*

A population comprised of 260 RIL family lines derived from the cross between 7235 and TM-1 using a bulk-selfing technique was employed to analyse the association between the *GhSusA1* DNA polymorphism and phenotypic effect on fiber qualities. There were different size fragments between TM-1 and 7235 amplified with *GhSuSA1* PCR primers. Single-marker association analysis revealed a significant correlation between the genomic level of this gene and fiber elongation, strength, micronaire and fiber maturity index, each explaining 1.11%, 4.52%, 6.12% and 9.76% of the phenotypic variance in 2002 and 2.29%, 5.47%, 6.76% and 11.35% of the phenotypic variance in 2003 (Table 1) in the 7235 and TM-1 RILs. These correlations of differential expressions of *GhSuSA1* between 7235 and TM-1 with corresponding changes in fiber elongation and secondary wall formation strongly suggest that GhSusA1 may play important roles in fiber development. However, this gene could be mapped on any linkage group constructed using this RIL because there are not enough loci in this intraspecific population (Shen *et al.*, 2007). Therefore, population derived from interspecific hybridization between Gh and Gb was used to map *GhSusA1* on chromosomes.

Table 1 Association analysis between fiber qualities and *GhSUSA1* polymorphism

Year	Fiber elongation (%)	Fiber strength (cN/tex)	Micronaire	Fiber maturity
2002				
7235*	5.77	36.29	4.18	0.90
TM-1†	5.63	34.39	3.84	0.88
F value	3.63	12.12	16.33	26.42
P	5.79 E-02	5.96 E-04	7.23E-05	5.18E-07
R^2	1.11%	4.52%	6.12%	9.76%
2003				
7235*	5.90	35.72	4.71	0.93
TM-1†	5.72	33.74	4.41	0.90
F value	5.51	13.59	17.04	30.11
P	1.97 E-02	2.83 E-04	5.10E-05	1.06E-07
R^2	2.29%	5.47%	6.76%	11.35%

RIL, recombinant inbred line.

*Data presented were the average of phenotypic values for 7235 genotypic RIL families.

†Data presented were the average of phenotypic values for TM-1 genotypic RIL families.

Southern blot analysis (Figure S4a) indicated that the *GhSusA1* gene is present in two copies in tetraploid cotton, with each subgenome most likely containing one copy. We isolated the genomic *GhSusA1* sequences from two subgenomes of *G. hirsutum* L acc. TM-1 (A-subgenome, GenBank accession number JF330272, and D-subgenome, GenBank accession number JF330273) and *Gossypium barbadense* cv. Hai7124 (A-subgenome, GenBank accession number JF330270, and D-subgenome, GenBank accession number JF330271), and their two diploid progenitors, *G. herbaceum* L. (A-genome, GenBank accession number HQ702188) and *Gossypium raimondii* Ulbrich (D-genome, GenBank accession number HQ702187). Analysis of the genomic DNA and cDNA sequence alignment of the *GhSusA1* orthologs from the four cotton species revealed that they share the same intron/exon structures, and their length variations were mainly caused by insertion/deletion events within the introns (Figure S5). Based on the 1.8% differences (0.3% in At genome and 1.5% in Dt genome) in genomic sequences between our two mapping parents, TM-1 and Hai7124 (Guo *et al.*, 2007), different subgenome polymorphisms were generated by using specific primers. It produced 112-bp (At) and 261-bp (Dt) fragment in TM-1, 108-bp (At) and 272-bp (Dt) fragment in Hai7124 by polyacrylamide gel electrophoresis. *GhSusA1* was mapped on homeologous chromosomes A8 (chro.8) and D8 (chro.24) (Figure S4b), which are densely populated with markers and quantitative trait loci (QTLs) associated with fiber strength, elongation and fineness from different groups of diverse cotton lines cross (Shen *et al.*, 2005, 2007; Qin *et al.*, 2008). As *GhSus1* was mapped on chromosome A5 (chro.5) (Zhu *et al.*, 2011), it is clear that *GhSusA1* is not a homologous and/or homeologous gene of *GhSus1*.

Expression and evolutionary analysis of the *GhSusA1* gene in tetraploid cotton

To understand the temporal–spatial distribution of *GhSusA1* gene transcripts, their expression was initially analysed *via* semi-quantitative RT-PCR in different tissues of TM-1. The results indicated that the *GhSusA1* transcripts were much lower in young leaves, sepals and fibers and relatively high in the roots, stamen and seeds of the plants at 10 DPA (Figure 2a). We further measured the expression level of *GhSusA1* in different phases of fiber development in *G. hirsutum* L. acc. TM-1 and *G. barbadense* L. cv. 7124 by real-time quantitative RT-PCR (qRT-PCR) using homeolog-specific primers for the given subgenome (see Experimental procedures). We found that the *GhSusA1* transcripts were expressed preferentially at the stages of fiber initiation and early elongation (0–8 DPA) and secondary cell wall thickening (20–23 DPA) (Figure 2b). Its expression pattern was different from that of *GhSus1*, which was previously found to be high throughout fiber development (Ruan *et al.*, 1997), and/or *GhSus2*, which was shown to be high in the secondary wall formation stage (Ruan *et al.*, Patent: WO 2008012058-A 31-JAN-2008). Furthermore, the abundance of *GhSusA1* transcripts in TM-1 was biased significantly towards the D-subgenome in most stages. However, the transcripts of the *GhSusA1* homeologs in Hai7124, which was a *G. barbadense* cultivar with longer fiber length than *G. hirsutum*, were equivalently expressed at the early stages, with only a D-subgenome bias observed from 17 to 23 DPA. This implied that the *GhSusA1* gene may play a functional role during fiber development and strong *GhSusA1* expression may contribute to higher fiber quality.

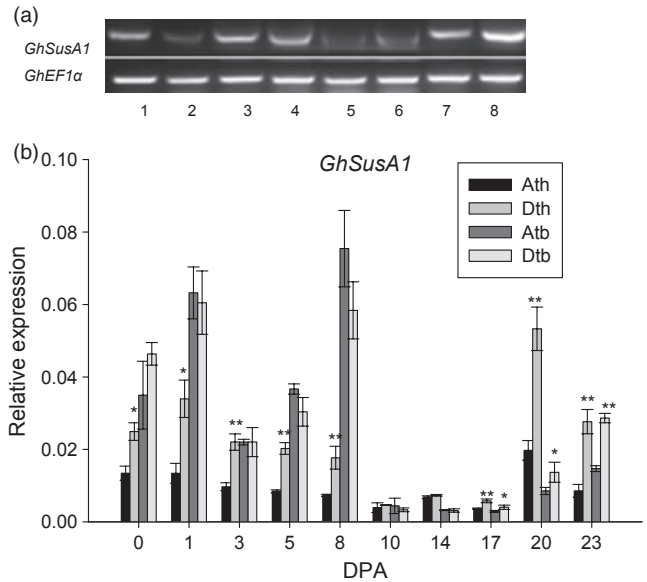

Figure 2 Analyses of *GhSUSA1* gene expression in cotton tissues. (a) Semi-quantitative RT-PCR analysis of expression of GhSUSA1 gene in cotton tissues. Expression level of *GhSUSA1* in cotton tissues, including root (8), stem (7), young leaf (6), sepal (5), petal (4), stamen (3), fiber (2) and ovule (1) at 11 days post-anthesis (DPA), was compared to *GhEF1α* expression activity. (b) Real-time RT-PCR analysis of *GhSusA1* with homeologous expression was conducted in ovules and fibers at different DPAs. The tissues before eight DPA were mixture of ovule and fiber, the others were fiber samples. Significant values were obtained by comparison between the two subgenomes. **P*-value ≤0.05, ***P*-value ≤0.01. Ath = A-subgenome of *Gossypium hirsutum* L. cv. TM-1, Dth = D-subgenome of *G. hirsutum* L. cv. TM-1, Atb = A-subgenome of *Gossypium barbadense* L. cv. 7124, Dtb = D-subgenome of *G. barbadense* L. cv. 7124. Vertical bars represented standard deviation (SD).

Seed yield and fiber quality are decreased in the antisense transgenic plants

From *GhSusA1* mapping, sequence comparison, phylogenetic relationship and expression pattern, it is clear that *GhSusA1* is a novel gene different from *GhSus1* and *2* in cotton. To investigate the function of *GhSusA1* in fiber development, its partial coding region (857 bp from ORF 280 to 1136) in reverse orientation was used to construct antisense *GhSusA1* transformation vectors driven by either a fiber-specific E6 (John and Crow, 1992) and/or a constitutive CaMV 35S promoters. The PCR product in the antisense orientation was subcloned into pBI121 and pBI121–E6 in which the CaMV 35S promoter was replaced by the GaE6 promoter (our unpublished data) by digestion with *Sma*I and *Sac*I to replace the *GUS* gene downstream of the CaMV 35S or E6 promoters, respectively. These two vectors were introduced into the cotton line W0 *via Agrobacterium*-mediated transformation (Figure S6) (Li *et al.*, 2009). The homozygous transgenic lines with 1–4 copies inserted were developed by pedigree selection and further used to conduct field tests for their yield and fiber quality performance. All of these lines containing either construct showed statistically significant decreases in fiber length compared with the corresponding control lines (Figure 3a,b).

Apart from a short-fiber phenotype, the vegetative growth and fruit size were also affected in those lines transformed with the vector driven by the 35S promoter to a greater extent. All of these transgenic antisense lines showed a decrease in seedling dry weight compared with both W0 and void-vector transformants (Figure 3c), which was not reported previously when *GhSus1* was suppressed (Ruan *et al.*, 2003). Although the difference in plant phenotype was not visibly evident until the boll stage between the transgenic and control lines, a statistically significant decline was still detected in seedling dry weight of three lines during the two true-leaves period, especially in Line 45 (Figure 3c). This result suggested that the vegetative growth was reduced from the early stage in the *35S-GhSusA1* suppressed lines. Furthermore, there was also a significant reduction in boll size and seed weight in all of the antisense lines driven by the 35S promoter (Figure 3d,e). Boll size and seed weight were reduced by 20% and 15%, respectively, in Line 45 compared with both wild-type and empty vector transgenic pBI121 plants (Figure 3f). However, there was no effect on boll size and seed weight in the *E6-GhSusA1* suppressed lines as expected because E6 promoter is active only in fibers. It is clear that the suppression of *GhSusA1* at the whole plant level caused a decline not only in fiber length, but also in biomass yield in cotton.

Overexpression of *GhSusA1* leads to increased cotton biomass yield

Further studies were conducted to explore whether overexpression of *GhSusA1* would enhance cotton development. To achieve this, the full-length ORF of *GhSusA1* was amplified by PCR, digested with enzymes and then ligated into the pBI121 vector between *Xba*I and *Bam*HI sites under the CaMV 35S promoter. The insertion of construct was confirmed by direct sequencing before transformation. The overexpression construct was introduced into W0 using the same method as that described for the antisense work. Because the gene suppression by both antisense suppression and co-suppression works through homology-dependent RNA degradation (Waterhouse *et al.*, 2001), a large number of primary transformants were generated to screen for *GhSusA1* overexpression lines. In this context, 14 lines in T3 were developed by pedigree selection

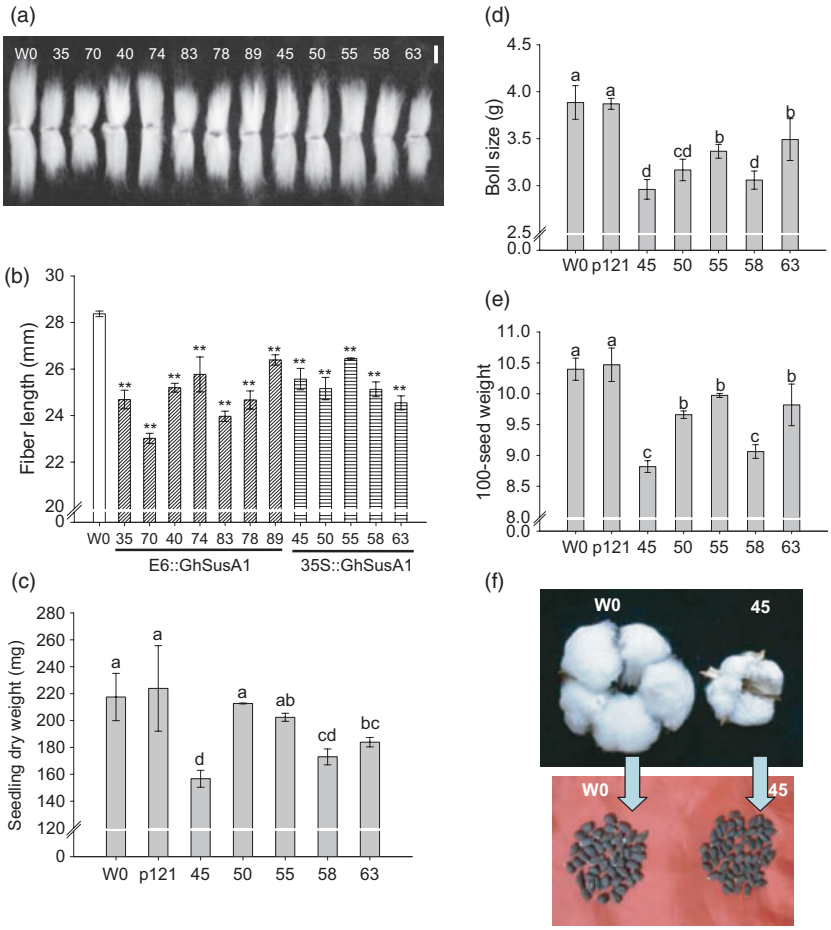

Figure 3 Seed yield and fiber quality are decreased in the *GhSusA1* antisense transgenic plants. (a) Short-fiber phenotypes of independent T3 transgenic homozygous lines. Bar = 10 mm. W0 was the recipient variety used for transformation, here served as a non-transgenic control. (b) The values of fiber length represent the average of triplicate experiments ± SD. Asterisks indicate significant difference from fiber length between transgenic lines and W0, ***P*-value ≤0.01 (Student's *t* test). (c) Statistical analysis of seedling biomass in different transgenic homozygous lines. Each value is the average of dry weight from 30 seedlings per independent line in two-leaf stage. Vertical bars represented standard deviation (SD). (d) Boll size data in transgenic lines compared with the control; the value is average weight of 60 mature bolls ± SD. (e) Effect of antisense suppression of *GhSusA1* on 100-seed weight. Mean ± SD was calculated from three replicates of per 100 seeds weight in transgenic line and W0. (f) Phenotyping reveals the small boll and seed size of transgenic Line 45 as compared to the control. Different letters in (c), (d) and (e) indicate significant difference at $P \leq 0.05$ according to randomization one-way ANOVA test. p121 in (c), (d) and (e) were the transformants with void-vector that did not contain GhSusA1 sequence.

assisted with kanamacin and transgene PCR selection. Southern hybridization analysis identified 12 independent transformants containing one to six copies of the inserted transgenes (Figure S7). Quantitative real-time PCR using gene-specific primers covering part of the 3′-UTR region revealed that there were seven lines with obvious co-suppression among the 12 different transformants. Therefore, the remaining five independent lines (16, 9, 17, 11 and 23) that exhibited higher *GhSusA1* expression (overexpressed lines) and two lines with the antisense *GhSusA1* gene (suppressed lines) were chosen for further detailed studies. The expression levels of *GhSusA1* in these five overexpressed lines, especially Line 9, were significantly higher than that in the W0 plants (Figure 4a). As seedling dry weights were decreased in all antisense *GhSusA1* lines (35S promoter) (Figure 3c), we examined the plant biomass at vegetative stage. The analysis showed that the biomass in the overexpressed plants appeared to increase from the early vegetative growth to the boll stage (Figure 4b). Importantly, the degree of biomass increment correlated with the level of *GhSusA1* transcript, with Line 9 displaying the highest expression level and biomass yield (Figure 4a vs. b) and growth vigour (Figure 4c). The enzyme assay (Figure 4d) showed that Sus activity in developing leaves was elevated in the overexpression lines. Thus, the increase in *GhSusA1* mRNA levels led to the enhancement of Sus enzyme activity in these five overexpression transgenic lines compared with wild-type and those antisense lines. At the same time, the carbohydrate content in leaves was measured by using anion-exchange high-performance liquid chromatography (HPLC). The analyses revealed a significant increase in hexose content in the leaves of the overexpressed lines (Figure 4e), as expected with the overexpression of genes encoding enzymes that degrade sucrose. Line 9 had the most significant increases in the fructose and glucose contents compared with the control, corresponding to the highest Sus activity in leaves, which led to a strong biomass accumulation and growth enhancement (Figure 4b,c). Overall, overexpressed lines had increased Sus activity corresponding to higher level of hexose in developing levels and plant biomass. These results demonstrate that

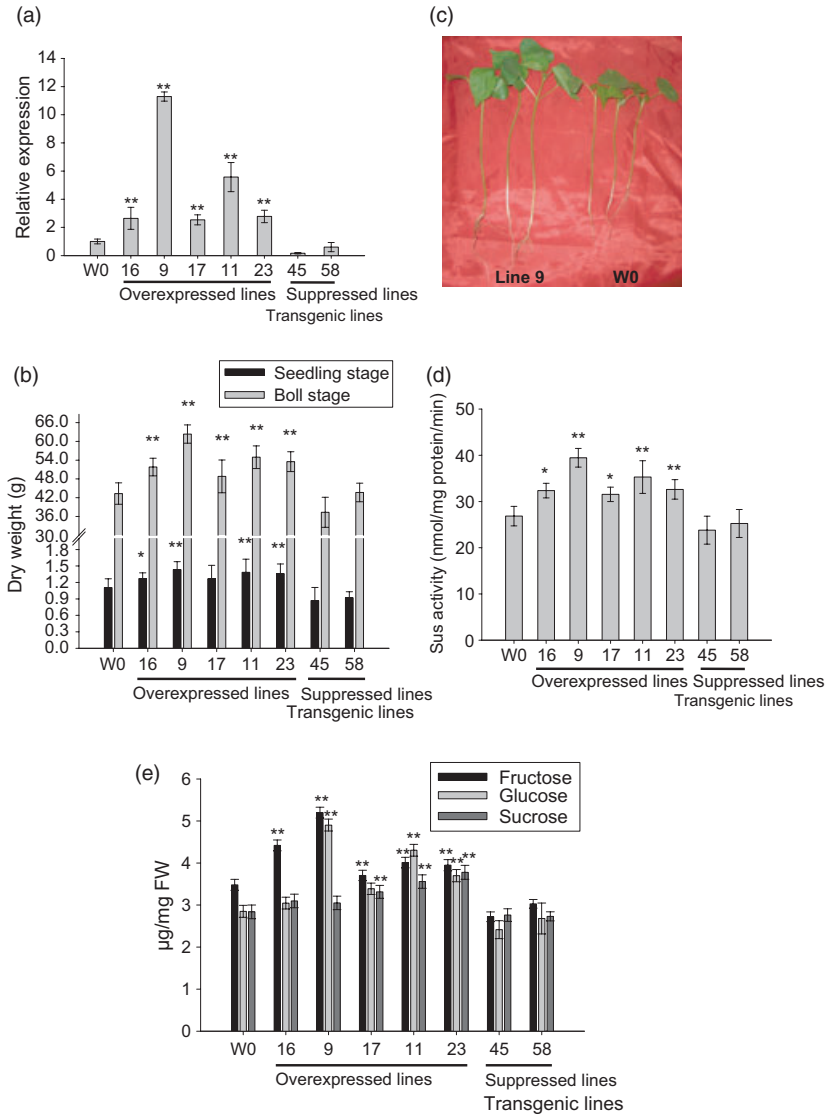

Figure 4 *GhSusA1* influenced plant biomass yield in cotton. (a) Transcript level of *GhSusA1* in sink leaves of transgenic cotton and wild type as measured by real-time qPCR. W0 represents non-transgenic plants, 'overexpressed lines' represents sense *GhSusA1* transformants and 'suppressed lines' represents antisense *GhSusA1* lines. Each value is the mean ± SD of three biological replicates. (b) Analysis of plant biomass during seedling stage (four true-leaves time) and blossom stage (time of the first flower bud appearance). Mean ± SD of plant dry weight was calculated from at least 15 plants per transgenic line at seedling time and at least 10 plants at boll stage. (c) Plant type of W0 and Line 9 under the same conditions showed more exuberant growth in Line 9 than wild type at the seedling time. (d) Sus activity in leaves of transgenic cotton plants. Vertical bars represented SD and calculated on three replicates. (e) Sugar content in leaves of wild-type and transgenic lines. Each value is the mean ± SD of at least three biological replicates. Asterisks indicate significant differences (Student's *t* test, $*P \leq 0.05$; $**P \leq 0.01$) between transgenic lines and W0.

GhSusA1 affects hexose metabolism and ultimately influence the biomass yield in cotton.

Fiber quality is improved in transgenic lines over-expressing *GhSUSA1*

Sus activity in 15 and 20 DPA fibers in the *GhSusA1* overexpressed Line 9 was increased by 15% and 22%, respectively (Figure 5a,b), which resulted in a significant increase in mature fiber length (Figure 5c,d). Furthermore, on average, 3.37%–12.25% increases in fiber strength were observed in the overexpressed lines planted in field in Hainan in 2009–2010 and Nanjing in 2010 (Table 2). Following elongation, the fibers began to thicken from about 20 DPA and wall thickness determines fiber strength (Lewis and Benedict, 1994). Transmission electron microscopic analyses revealed that the fiber cell wall in the overexpressed Line 9 was much thicker than that of the wild type, whereas the *GhSusA1*-suppressed fiber appeared thinner than the wild type at 20 DPA (Figure 5e–g). The phenomenon is correlated positively with *GhSusA1* mRNA levels (Figure 5a) and enzyme activity (Figure 5b). This accelerated thickening in Line 9 continued until 25 DPA, the peak of secondary wall synthesis (Figure 5h–j), and seemed to result in much thicker fibers than those of the wild type (Figure 5i vs. h). Such a difference in cell wall thickness was further confirmed by statistical analysis of the mature fibers. The thickening rate in *GhSusA1*-overexpressed fibers was observed in the electron

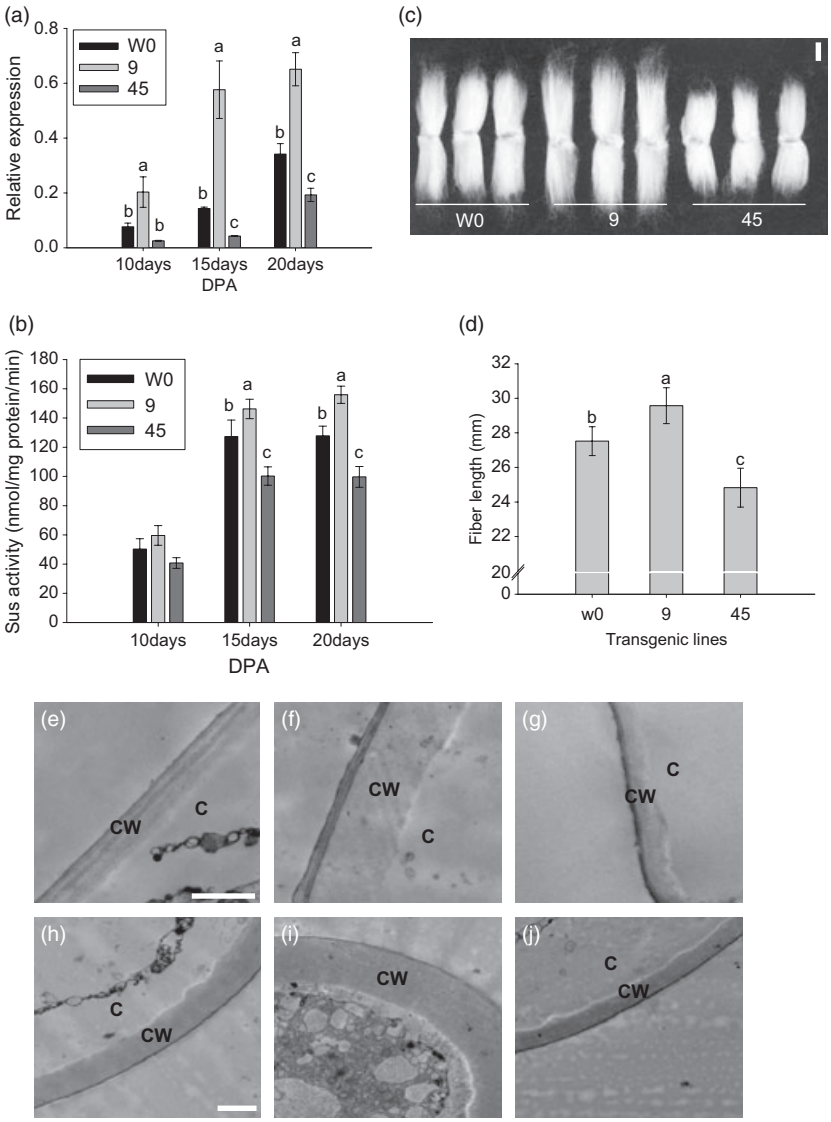

Figure 5 Sus transcript level and enzyme activity, and phenotype of fibers from transgenic cotton transformed with sense *GhSusA1* (Line 9) and antisense *GhSusA1* (Line 45). (a) Relative values of *GhSusA1* expression in fibers at 10, 15 and 20 days post-anthesis (DPA) of transgenic lines (9 and 45) and wild type (W0). Vertical bars represent SD, $n = 3$. (b) Sus activity in cotton fibers at 10, 15 and 20 DPA. Vertical bars represent SD, $n = 3$. (c) Photograph of cotton fibers in wild-type (W0) and transgenic plant (Line 9 and Line 45). Short Bar = 10 mm. (d) Average fiber length of wild-type and transgenic plants measured from T3 generations. Values are mean ± SD, calculated from 10 replicates per transgenic line. (e)–(j) Cell wall morphology of the cotton fibers. Transmission electron microscopy analysis of cotton fibers was conducted in W0 (e and h), Line 9 (f and i) and Line 45 (g and j). Thickness of secondary cell wall was showed at 20 DPA in (e), (f) and (g), at 25 DPA in (h), (i) and (j). CW, cell wall; C, cytoplasm; bars = 1 μm in (e) as the same as scale for (f) and (g), and 2 μm in (h) as the same as scale for (i) and (j). Significant values were obtained by comparison among different lines. The different letters in (a), (b) and (d) indicate significant difference in the specified stage at $P \leq 0.05$ according to randomization one-way ANOVA test.

microscopy images to be about 0.34 ± 0.02 μm/day, much higher than the rate in the wild type (0.20 ± 0.01 μm/day) from 20 to 25 DPA. By contrast, this rate was only 0.14 ± 0.01 μm/day in the *GhSusA1*-suppressed fibers. These different thickening rates may reflect the different rates of cellulose deposition, which could be the most important factor contributing to fiber strength (Lewis and Benedict, 1994). There was no significant change of *GhSus1* in the fibers at the transcriptional level relative to controls (Figure S8), suggesting that the improvement in fiber qualities described above was not caused by *GhSus1*, although it was expressed and influenced fiber initiation and elongation (Ruan *et al.*, 2003). At the same time, we also detected the expression of *GhSus2* and found it was always low during this time, with the same as that described in the patent in 2008 (Ruan *et al.*, Patent: WO 2008012058-A 31-JAN-2008) and was not affected by overexpression or suppression of *GhSusA1*.

Discussion

All plant species appear to have a *Sus* gene family. In Arabidopsis, there are six non-allele Sus genes in total, with different expression patterns in different tissues (Baud *et al.*, 2004). At least three different types of Sus genes have been found in

Table 2 Fiber strength in the overexpressed lines planted in field in Hainan in 2009–2010 and Nanjing in 2010

Transgenic lines number	Strength (cN/tex) (2009-2010)	Strength (cN/tex) (2010)
W0	25.25 ± 0.05	27.50 ± 0.36
16	27.05 ± 0.65**	29.37 ± 1.43*
9	28.00 ± 0.50**	30.87 ± 1.23**
17	26.10 ± 0.70	30.43 ± 1.44**
11	26.75 ± 0.05**	30.47 ± 0.65**
23	26.15 ± 0.65	29.30 ± 0.17*

Each value is the mean ± SD of three replicates. Significant values were obtained by comparison between W0 and *GhSusA1*-overexpressed lines, **P*-value ≤0.05, ***P*-value ≤0.01.

cotton, which differ in intron structures (Figure 1c). The homology among the three cotton Sus genes is far less than the same type of Sus from different species. The three cotton Sus genes belong to different evolutionary branches (Figure S2). *GhSusA1* is different from the previously reported Sus genes (Ruan *et al.*, 2003 and Patent: WO 2008012058-A 31-JAN-2008) both in structure and evolution (Figures 1c and S2). Repression of *GhSus1* gene blocks the initial development of cotton fiber as well as seed development (Ruan *et al.*, 2003). Inhibition of GhSus2 also reduces the accumulation of cellulose (Ruan *et al.*, 2008; Patent: WO 2008012058-A 31-JAN-2008). However, the result of Q-PCR in our research showed that both suppression and overexpression of GhSusA1 had little effect on the transcript level of GhSus1 (Figure S8). And they had no effect on the GhSus2 either, which is not surprising because the expression of GhSus2 was very low in wild type. The intron structure of GhSusA1 gene is more complex relative to the other two genes. Although the exact function of the intron remains unknown, intron structure or position may be related to fine tuning of gene regulation in plant (Richmond and Somerville, 2000). As a member of Sus family genes in cotton, it is possible that GhSusA1 plays a collaborative or complemental role with other Sus isozyme. For example, if one Sus gene is mutated because of environmental stress, another isoform may serve as a backup to sustain the Sus activity required for sucrose metabolism and normal development. Our transgenic results suggest GhSusA1 partly overlaps with GhSus1 in function, but has additional role in vegetative growth.

The expression pattern of *GhSusA1* indicates its possible role in the early stage of fiber initiation and the secondary wall formation stage. The expression of Sus at these stages may be physiologically important because sucrose moves to fiber symplasmically through plasmodesmata at these two stages (Ruan *et al.*, 2001), where Sus plays a pivotal role in degrading the incoming sucrose (Ruan *et al.*, 2005, 2010). The fact that *GhSusA1* was expressed at a significantly higher level in Hai7124 than in TM-1 at the initiation and early elongation period suggests a role of GhSusA1 in fiber elongation.

Sus has been shown to play a major role in sucrose metabolism in a wide range of sink tissues. Cotton fibers are highly specialized single cells, characterized with rapid and extensive cell elongation and secondary cell wall synthesis. These two processes determine two important fiber traits, length and strength, respectively. The role of Sus in cotton development was reported previously (Ruan *et al.*, 2003, 2008), and many QTLs have been tagged (Shen *et al.*, 2005; Qin *et al.*, 2008). However, the molecular mechanisms controlling cotton fiber development remain largely unknown. In this study, we isolated a novel *Sus* gene, *GhSusA1*, from a super quality fiber *G. hirsutum* line 7235 and found that its expression level was highly correlated with fiber qualities in RILs derived from hybridization between 7235 and a standard line TM-1. Duplicate *GhSusA1* genes were mapped on A8 (chro.8) and D8 (chro.24) chromosomes, and its expression is significantly higher in D8 (chro.24) than in A8 (chro.8) from 17 DPA in both TM-1 and Hai7124 (Figure 2b). The chromosome D8 (chro.24) was reported to contain many QTLs for fiber qualities (Shen *et al.*, 2007; Qin *et al.*, 2008). Thus, *GhSusA1* may contribute to fiber quality formation. Indeed, association analyses between *GhSusA1* and fiber quality in RILs of 7235× TM-1 show that this gene has specific relationships with fiber elongation, strength, micronaire and maturity, contributing to the phenotype at different rates from 2.29% to 11.35% in our field experiments conducted in 2003 (Table 1). Importantly, overexpression of GhSusA1 in transgenic cotton significantly increased in fiber length in Line 9 (Figure 5c), which was in contrast to the short-fiber phenotype observed in the antisense transgenic lines (Figure 3a). Moreover, along with the increase in fiber length, the fiber strength in five overexpressed lines also mildly improved. These observations demonstrate the critical role of GhSuSA1 in cotton fiber elongation and cellulose biosynthesis.

Increase in fiber length may be attributed to increased turgor pressure as reported previously (Ruan *et al.*, 2000, 2001). By degrading sucrose into UDP-glucose and fructose, Sus could generate osmotic potential, hence turgor pressure, for fiber elongation and provide substrate, UDP-glucose, for cellulose biosynthesis (Ruan *et al.*, 2003). The enhancement in fiber strength could be attributable to the observed higher rate of secondary wall thickening. Here, enhanced wall thickening in a overexpressed Line 9 (Figure 5e–j) was correlated with the increase in *GhSusA1* expression level and Sus activity at 20 DPA (Figure 5a,b). Overexpression of Sus could enhance the cell wall integrity by increasing the supply of UDP-glucose, the substrate for the synthesis of both cellulose and many non-cellulose cell wall compounds (Buchala, 1999). Therefore, Sus is a key enzyme in secondary wall formation of cotton fiber, and its change in activity was directly associated with fiber quality in the transgenic plants as well as in the population of the RIL.

Besides the positive impact on fiber quality by overexpression of *GhSusA1*, another new observation was the effect on plant biomass by altering *GhSusA1* expression. In the present study, suppression of *GhSusA1* had a more pronounced effect on seedling dry weight and corresponding boll and seed weight (Figure 3c–f) than on fiber quality changes, indicating a role of this gene in controlling sink strength. The down-regulation of Sus has been shown to influence biomass in other plants. For example, carrot plants with antisense-mediated suppression of Sus were small and had few leaves (Tang and Sturm, 1999). Similarly, tomatoes with inhibited Sus had a reduced fruit size and compromised sucrose unloading capacity (D'Aoust *et al.*, 1999). Reducing Sus expression and activity could decrease sucrose unloading into sinks, and thus availability of hexose pools for cellular metabolism and biosynthesis, leading to low plant biomass. As suppression of *GhSusA1* expression did not affect floral formation and flowering from our research, the decline in fiber and seed size may be ascribed in part to the negative effects on vegetative growth.

In contrast to the situation in the suppressed lines, the plants with overexpressed *GhSusA1* showed increased biomass during the seedling and boll stages, demonstrating that overexpressing this gene can elevate biomass yield in cotton. A similar influence in biomass has been reported previously. In poplar, enhanced growth was observed in some lines transformed with a modified *Sus* gene from mung bean under the control of the 35S promoter (Konishi *et al.*, 2004). In tobacco, overexpression of *GhSus1* from cotton also resulted in increased biomass and metabolism (Coleman *et al.*, 2006).

In conclusion, our genetic and transgenic analyses identified a novel sucrose synthase gene, *GhSusA1*. Overexpression of this gene in cotton increased plant biomass and cotton fiber yield and fiber quality. This study provides evidence for a direct connection between sucrose metabolism and cotton biomass or fiber traits through the activity of Sus. The discoveries have significant implications in improving cotton productivity and fiber quality.

Experimental procedures

Plant materials and growth conditions

Cotton (*G. hirsutum* L. acc 7235 and TM-1) plants were grown in a greenhouse at the Nanjing Agricultural University. TM-1 is the genetic standard accession of *G. hirsutum* (Kohel *et al.*, 1970), whereas 7235, developed in China, is an elite fiber strain of *G. hirsutum* (Qian *et al.*, 1992). Tissues for DNA and RNA extraction were derived from cotton plants grown under standard field conditions. Developing ovules were excised from the flower buds or bolls on specified days before or post-anthesis relative to the day of anthesis (0 DPA). Cotton (*G. hirsutum* cv W0) seeds were surface-sterilized with 70% ethanol for 30–60 s and 10% H_2O_2 for 60–120 min, followed by washing with sterile water. Sterilized seeds were germinated on half-strength MS medium under a 16-h light/8-h dark cycle at 26 °C. Cotyledons and hypocotyls were cut from sterile seedlings as explants for transformation.

Molecular cloning and sequence analysis of *GhSusA1*

The *GhSusA1* clone from the 7235 cDNA library (preserved in our institution) was isolated by amplifying the full-length sequence using TAIL (Liu and Whittier, 1995) (Table S1) and RACE techniques. PCR was performed on the DNA and cDNA of 7235, and the partial fragment was used as the template for designing primers. The genomic sequence was spliced and analysed using DNAstar software. Full-length *GhSusA1* genomic DNA was amplified in *G. herbaceum* L. and *G. raimondii* Ulbrich using the sense primer g-full-F (5′-GATACGAAGTTT AACGATGGCG-3′) and the antisense primer g-full-R (5′-AAA ATAGCAGCAAGCGAGACC-3′) for obtaining the A-genome and D-genome sequences, and sub-genome sequences of tetraploid cottons. The full-length *GhSusA1* cDNA sequence was amplified by PCR from the cDNA of 7235 fibers using Pyrobest™ DNA Polymerase (TaKaRa, Dalian, China) and the primers c-full-F (5′-CACGGACCGATTGCCTCA-3′) and c-full-R (5′-TCATCCTCGCAATCAACCAGA-3′). Amino acid sequences encoded by the *Sus* genes from different plants species were chosen using Blast software from NCBI. Sequence alignments were carried out with the ClustalX software (version 1.81) (Thompson *et al.*, 1997). The export and editing of these sequences were conducted with the GeneDoc software (Nicholas and Nicholas, 1997).

DNA gel blot analysis

TM-1 genomic DNA was isolated following the protocol by Paterson *et al.* (1993) and digested with *Eco*RI, *Hin*dIII, *Xba*I and *Bam*HI. The DNA was then separated on 0.8% agarose gels and transferred onto Hybond N+ nylon membranes by capillary blotting. The DIG-labelled partial *GhSusA1* cDNA (the same as the antisense transgenic fragment) was used as a probe for the hybridization experiment according to the instructions described in the Dig High Prime DNA Labeling and Detection Starter Kit I (Roche, Basel, Switzerland). For Southern blot analysis of transgenic antisense and sense lines, the genomic DNA of transgenic plants was digested with *Eco*RI that cleaves the T-DNA at a unique site. A 750-bp fragment of *NPTII* coding region amplified using the primers *NPTII*-F (5′-GAGGCTATTCGGCTATG-ACTG-3′) and *NPTII*-R (5′-TAGAAGGCGATGCGCTGCGA-3′) was used as a hybridization probe.

RT- and real-time PCR analyses

Total cellular RNA was extracted using the CTAB-sour phenol extraction method (Jiang and Zhang, 2003). Each RNA sample was treated with DNase I after the extraction to remove residual DNA. The expression of the *GhSusA1* in cotton tissues was analysed by semi-quantitative and real-time RT-PCR. A cotton elongation factor (*EF1α*) gene was used as a standard control in the RT-PCR. Total RNA samples (2 μg per reaction) from different tissues were reversely transcribed into cDNAs by AMV reverse transcriptase. The cDNAs were then used as templates in RT-PCR with (i) *GhSusA1* gene-specific primers that annealed to part of the 3′-UTR region (SUSA-RT-F: 5′-TGAGT-GATCGGTCAAAGCC-3′; SUSA-RT-R: 5′-CGGTCTGGTCA GGGTGGTA-3′) and (ii) *GhEF1α* gene-specific primers (EF1α-RT-F: 5′-AGACCACCAAGTACTACTGCAC-3′; EF1α-RT-R: 5′-CCACCAATCTTGTACACATCC-3′). The homeolog-specific RT-PCR primers were designed by SNP (At: 5′-TTCTACGAA GCCTTTGGACTTAGA-3′, Dt: 5′-ACGAAGCCTTTGGACTCA CG-3′, and the same reverse primer: 5′-TTAAGCCCTCCATC-AGATATTTTAGT-3′). The homeolog-specific primers were designed and applied based on sequence differences between duplicated loci. Specificity of those primers was detected by PCR amplification of genomic DNA from *G. herbaceum* (A-genome), *G. raimondii* (D-genome), TM-1 and Hai7124. Primers specific for the A-genome resulted in amplifications from *G. herbaceum*, TM-1 and Hai7124, but not from *G. raimondii*, which indicated that the amplification products of TM-1 and Hai7124 were from their A-subgenomes. Conversely, the D-genome-specific primers only amplified from *G. raimondii* and D-subgenome of the allopolyploid. The real-time PCR was performed using the light cycler fast start DNA Master SYBR Green I kit (Roche) according to the manufacturer's instructions. The relative expression levels of *GhSusA1* were calculated using the equation $Y = 2^{-\Delta C_t}$ (where C_t is the value at which the amplification signal crosses the positive threshold, and ΔC_t is the difference in the C_t of the target *GhSusA1* and the control *GhEF1α* signals; i.e., $\Delta C_t = C_{tGhSusA1} - C_{tGhEF1\alpha}$). The detailed information of real-time PCR was submitted in Appendix S1 with MIQE checklist. We chose different plants from the same lines planted at the same time and grown in the same conditions to mix and then extract RNA to conduct real-time PCR with three biological reps.

Heterologous expression of GhSusA1 in *Escherichia coli*

To express the putative mature protein in *E. coli*, the open reading frame of *GhSusA1* was amplified by PCR using the primers SUSA-pet-F (5′-TATCGGGATCCATGGCGAATCCTAAGCTC GGTCG-3′) and SUSA-pet-R (5′-ATAACGCGTCGACATCATCAC TGGCCAAGGGAAC-3′). The PCR-amplified fragments were then digested with *Bam*HI and *Sal*I and cloned into the *Bam*HI plus *Sal*I-digested pET-30a (+) vector (Qiagen, Valencia, CA), resulting in a translational fusion of the protein with six His residues at the N and C termini. Cultures of *E. coli* carrying pET-30a (+) –*GhSusA1* were grown at 30 °C in LB medium with 50 mg/L kanamycin to an OD_{600} of 0.3. Expression of the fusion protein was induced by addition of 0.6 mM isopropylthio-β-galactoside (IPTG), and the culture was allowed to grow at 30 °C for 12 h. Protein was purified using a Ni–NTA-agarose column (Qiagen, Hilden, Germany) according to the manufacturer's instructions. Crude and purified proteins were used for enzyme assays and SDS–polyacrylamide gel electrophoresis (SDS–PAGE) analysis.

Molecular mapping and association analysis of the *GhSusA1* gene with fiber quality phenotypes

Because *GhSusA1* cannot be mapped with the (7235× TM-1) RIL populations (Shen *et al.*, 2007), we mapped this gene using the BC_1 ([(TM-1 × Hai7124) × TM-1]) interspecific mapping population from our laboratory. The primer pairs At F (5′-CTGC CTCTTGAAATCTGGTT-3′), At R (5′-AAAATAGCAGCAAGCG AGAC-3′) and Dt F (5′-GTTTGACGTATGGCCTTTTC-3′), Dt R (5′-CTGGAATACCCTGCAACTCT-3′), which amplified regions with different subgenome polymorphisms between TM-1 and Hai7124, were used to survey 138 individuals of the BC_1 mapping population. The polymorphic loci were integrated in our backbone map (Guo *et al.*, 2008) by Join Map3.0 (Van Ooijen and Voorrips, 2001). The population used to conduct association analysis of DNA polymorphism and phenotypic effect for fiber quality was comprised of 260 RILs that were generated from an F_2 population of an Upland cotton cross (7235× TM-1). The $F_{6:8}$ RILs derived from the cross between two Upland cotton germplasm lines, using bulk-selfing technique. The method of 'single grain spread' was used to generate offspring. Based on amplification polymorphisms of the genomic *GhSusA1* sequences in 7235 and TM-1, the primers Dt F and R were used to detect the polymorphisms in the two parents. Phenotypic data were collected at Nanjing and Guanyun County of Jiangsu province in 2002 and 2003 (Shen *et al.*, 2007). Statistical analyses of the two genotypes were performed using the *F* test and regression analysis. Correlation coefficients were calculated based on means of RILs at two locations in 2002 and 2003 using SAS 6.0 program (SAS Institute Inc., Cary, NC, USA). The coefficient of genetic determination R^2 was used to explain the phenotypic variation.

Construction of the plant expression *GhSusA1* vector

The 857-bp partial coding region of *GhSusA1* (from ORF 280 to 1136) was digested from the cloning vector (pGEM T-easy, Promega) by *Eco*RV and *Sac*I and then subcloned into the *Sma*I- and *Sac*I-digested pBI121 and pBI121–E6 vectors (a pBI121 derivative where the CaMV 35S promoter was replaced by the *GaE6* promoter) in the antisense orientation, which replaced the *GUS* gene downstream of the CaMV 35S promoter and E6 promoter, respectively. To construct the sense *GhSusA1* vector, the full-length ORF of *GhSusA1* was amplified by PCR (primers, sense vector-F: 5′-GCTCTAGACGGACCGATTGCCTCA-3′; sense vector-R: 5′-CGGGATCCATCATCACTGGCCAAGGG-3′) and the sites *Xba*I and *Bam*HI were introduced to the ends of the target gene. *GhSusA1* was inserted into the pBI121 vector between the *Xba*I and *Bam*HI sites. The constructs were sequenced to ensure that they did not contain any PCR or cloning errors and used for cotton transformation.

Cotton transformation and transgenic plant selection

Cotton transformation was performed as previously described (Li *et al.*, 2009). The constructs were introduced into the *Agrobacterium* strain LBA4404 for transformation. Cotyledon and hypocotyls explants from *G. hirsutum* cv W0 were transformed using *Agrobacterium*-mediated transformation and selections were performed on kanamycin sulphate-containing medium. After about 10 months of *in vitro* culture and selection, the putative transgenic plants were transferred to a glasshouse for further screening and analysis. Homozygosity of transgenic plants was determined by segregation analyses based on the presence or absence of kanamycin selection marker, PCR analysis for the *GhSusA1* transgene, and further confirmed by DNA gel blot and real-time PCR. *NPTII* specific primers for PCR detections were as described above. A second set of primers was used to determine the presence of the promoters and the *GhSusA1* transgene: 35S-F (5′-CACAATCCCACTATCCTTCG-3′), E6-F (5′-GGACCACAATCATCACCAT-3′) and SUSA-a (5′-ATA-CTTCGGGCAAGCAAA-3′) for the antisense and SUSA-b (5′-CCACCAGTGTCGGGCAAA-3′) for the sense constructs.

The homozygous transgenic lines were developed by pedigree selection and further used to conduct field tests for their yield and fiber quality performance in Hainan in 2009–2010 and Nanjing in 2010. Fiber samples of transgenic lines and W_0 were tested at the Supervision, Inspection, and Test Center of Cotton Quality, Ministry of Agriculture in China. Fiber quality traits tested mainly included fiber length and fiber strength. Biomass analysis was carried out by the determination of the dry weight of the cotton seedlings.

Enzyme activity assays

Samples (approximately 0.5 g each) were ground to a fine powder in liquid N_2. The grinding continued for 5 min in cold extraction buffer (3 : 1, v/w) containing 25 mM Hepes–KOH (pH 7.3), 5 mM EDTA, 1 mM DTT, 0.1% soluble PVP, 20 mM β-mercaptoethanol, 1 mM PMSF and 0.01 mM leupeptin. The homogenate was centrifuged at 10 000 ***g*** for 5 min at 4 °C. Sus activity was measured as the rate of sucrose cleavage (Chourey, 1981). The resultant reducing sugars were estimated according to the method of Ruan *et al.* (2003). The control for the Sus activity assay lacked UDP.

Soluble carbohydrate analysis

The method of carbohydrate deposition was performed as Ruan *et al.* (2003). Approximately 0.2 g of fiber or seed tissue was ground in liquid nitrogen. The samples were extracted with 4 mL of preheated 80% ethanol for 5 min at 80 °C. After cooling, they were centrifuged at 12 000 ***g*** for 10 min. The supernatants were collected, whereas the pellets were resuspended in 2 mL of 50% ethanol and respun as described above. The resulting pellet was re-extracted with 2 mL of water and re-centrifuged. The total 8 mL of supernatant collected was mixed with an equal volume of chloroform and shaken vigorously. The aqueous phase

was collected, dried in a vacuum and redissolved in 1 mL of water. The sucrose, glucose and fructose contents were measured and analysed using anion-exchange HPLC system (Agilent 1100 series; Agilent technologies, Inc., Santa Clara, CA, USA), which consisted of a G1311A pump and a G1362A refraction index detector. Sugar compounds were separated on a Sugar-D column (4.6 × 250 mm, Nacalai Tesque Inc., Japan), using acetonitrile/water (75 : 25, v/v) as the mobile phase at the flow rate of 1.0 mL/min. The injection volume was 40 μL. Quantification of each sugar was accomplished by comparing the peak areas of the samples with those of the standard solutions.

Electron microscopy analyses of cell wall thickness and measurement on tissue sections

For transmission electron microscopy analysis, cotton fibers (at 20 and 25 DPA) were fixed for 12 h in 2.5% glutaraldehyde. After three washes, the fibers were incubated in 1% osmium tetraoxide for 4 h at room temperature. The specimens were washed three times and dehydrated in a series of ethanol (50%–90%) and epoxy ethane and then embedded in Epon resin for 48 h at 60 °C. The middle parts of the fibers were cut into thin sections (50 nm) on an LKBIII Ultratome, stained with uranyl acetate followed by lead citrate, then examined and photographed under a Hitachi H-7650 transmission electron microscope.

Acknowledgements

This programme was financially supported in part by the State Key Basic Research and Development Plan of China (2011CB109300), National Science Foundation in China (30730067), National Research and Development Project of Transgenic Crops of China (2010ZX08009-003) and the Program for 111 project (B08025).

GenBank accession numbers

Genome sequence of *GhSusA1* from 7235 line (GenBank accession number HQ702186).
Transcription sequence of *GhSusA1* from 7235 line (GenBank accession number HQ702185).
A-subgenome sequence of *GhSusA1* come from *G. hirsutum* L acc. TM-1 (GenBank accession number JF330272).
D-subgenome sequence of *GhSusA1* come from *G. hirsutum* L acc. TM-1 (GenBank accession number JF330273).
A-subgenome sequence of *GhSusA1* come from *G. barbadense* cv. Hai7124 (GenBank accession number JF330270).
D-subgenome sequence of *GhSusA1* come from *G. barbadense* cv. Hai7124 (GenBank accession number JF330271).
Genome sequence of *GhSusA1* from diploid *G. herbaceum* L. (A-genome, GenBank accession number HQ702188).
Genome sequence of *GhSusA1* from diploid *G. raimondii* Ulbrich (D-genome, GenBank accession number HQ702187).

References

Amor, Y., Haigler, C.H., Johnson, S., Wainscott, M. and Delmer, D.P. (1995) A membrane-associated form of sucrose synthase and its potential role in synthesis of cellulose and callose in plants. *Proc. Natl Acad. Sci. USA*, **92**, 9353–9357.

Baud, S., Vaultier, M.N. and Rochat, C. (2004) Structure and expression profile of the sucrose synthase multigene family in Arabidopsis. *J. Exp. Bot.* **55**, 397–409.

Buchala, A.J. (1999) Noncellulosic carbohydrates in cotton fibers. In *Cotton Fibres Developmental Biology, Quality Improvement, and Textile Processing* (Basra, A.S., ed.), pp. 113–136. New York: Haworth Press.

Buckeridge, M.S., Vergara, C.E. and Carpita, N.C. (1999) The mechanism of synthesis of a mixed-linkage (1–3), (1–4) β-D-glucan in maize: evidence for multiple sites of glucosyl transfer in the synthase complex. *Plant Physiol.* **120**, 1105–1116.

Chourey, P.S. (1981) Genetic control of sucrose synthetase in maize endosperm. *Mol. Gen. Genet.* **184**, 372–376.

Chourey, P.S. and Nelson, O.E. (1979) Interallelic complementation at the sh locus in maize at the enzyme level. *Genetics*, **91**, 317–325.

Christine, A.R.W. and Sean, M. (1998) Activity of the yeast MNN1 α-1, 3-mannosyltransferase requires a motif conserved in many other families of glycosylransferases. *Proc. Natl Acad. Sci. USA*, **95**, 7945–7950.

Coleman, H.D., Ellis, D.D., Gilbert, M. and Mansfield, S.D. (2006) Up-regulation of sucrose synthase and UDP-glucose pyrophosphorylase impacts plant growth and metabolism. *Plant Biotech. J.* **4**, 87–101.

D'Aoust, M.A., Yelle, S. and Nguyen-Quoc, B. (1999) Antisense inhibition of tomato fruit sucrose synthase decreases fruit setting and the sucrose unloading capacity of young fruit. *Plant Cell*, **11**, 2407–2418.

Delmer, D.P. (1999) Cellulose biosynthesis: exciting times for a difficult field of study. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **50**, 245–276.

Geigenberger, P. and Stitt, M. (1993) Sucrose synthase catalyses a readily reversible reaction in vivo in developing potato tubers and other plant tissues. *Planta*, **189**, 329–339.

Guo, W.Z., Cai, C.P., Wang, C.B., Han, Z.G., Song, X.L., Wang, K., Niu, X.W., Wang, C., Lu, K.Y., Shi, B. and Zhang, T.Z. (2007) A microsatellite-based, gene-rich linkage map reveals genome structure, function and evolution in Gossypium. *Genetics*, **176**, 527–541.

Guo, W.Z., Cai, C.P., Wang, C.B., Zhao, L., Wang, L. and Zhang, T.Z. (2008) A preliminary analysis of genome structure and composition in *Gossypium hirsutum*. *BMC Genomics*, **9**, 314 (doi: 10.1186/1471-2164-9-314).

He, Y.J., Guo, W.Z., Shen, X.L. and Zhang, T.Z. (2008) Molecular cloning and characterization of a cytosolic glutamine synthetase gene, a fiber strength-associated gene in cotton. *Planta*, **228**, 473–483.

Jiang, J.X. and Zhang, T.Z. (2003) Extraction of total RNA in cotton tissues with CTAB-acidic Phenolic method. *Cotton Sci.* **15**, 166–167 (in Chinese with English abstract).

John, M.E. and Crow, L.J. (1992) Gene expression in cotton (*Gossypium hirsutum* L.) fiber: cloning of the mRNAs. *Proc. Natl Acad. Sci. USA*, **89**, 5769–5773.

Kohel, R.J., Richmond, T.R. and Lewis, C.F. (1970) Texas marker-1 description of genetic standard for *Gossypium hirsutum* L. *Crop Sci.* **10**, 670–671.

Komatsu, A., Moriguchi, T., Koyama, K., Omura, M. and Akihama, T. (2002) Analysis of sucrose synthase genes in citrus suggests different roles and phylogenetic relationships. *J. Exp. Bot.* **53**, 61–71.

Konishi, T., Ohmiya, Y. and Hayashi, T. (2004) Evidence that sucrose loaded into the phloem of a poplar leaf is used directly by sucrose synthase associated with various β-glucan synthases in the stem. *Plant Physiol.* **134**, 1146–1152.

Lewis, H.L. and Benedict, C.R. (1994) Development of cotton fiber strength. In *Biochemistry of Cotton* (Benedict, C.R. ed.), pp. 121–123. Galveston, TX: Cotton Incorporated.

Li, F.F., Wu, S.J., Chen, T.Z., Zhang, J., Wang, H.H., Guo, W.Z. and Zhang, T.Z. (2009) *Agrobacterium*-mediated co-transformation of multiple genes in Upland cotton. *Plant Cell Tissue Organ Cult.*, **97**, 225–235.

Liu, Y.G. and Whittier, R.F. (1995) Thermal asymmetric interlaced PCR: automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking. *Genomics*, **25**, 674–681.

Nicholas, K.B. and Nicholas, H.B. (1997) GeneDoc: a tool for editing and annotating multiple sequence alignments. Distributed by the author (http://www.cris.com/ketchup/genedoc.shtml).

Paterson, A.H., Brubaker, C. and Wendel, J.F. (1993) A rapid method for extraction of cotton (*Gossypium* spp.) genomic DNA suitable for RFLP or PCR analysis. *Plant Mol. Bio. Rep.* **11**, 122–127.

Pugh, D.A., Offler, C.E., Talbotc, M.J. and Ruan, Y.L. (2010) Evidence for the role of transfer cells in the evolutionary increase in seed and fiber biomass yield in cotton. *Mol. Plant*, **3**, 1075–1086.

Qian, S.Y., Huang, J.Q., Peng, Y.J., Zhou, B.L., Ying, M.C., Shen, D.Z., Liu, G.L., Hu, T.X., Xu, Y.J., Gu, L.M., Ni, W.C. and Chen, S. (1992) Studies on the hybrid of *G. hirsutum* L. and *G. anomalum* Wawr. & Peyr. and application in breeding. *Sci. Agric. Sin.* **25**, 44–51 (in Chinese).

Qin, H.D., Guo, W.Z., Zhang, Y.M. and Zhang, T.Z. (2008) QTL mapping of yield and fiber traits based on a four-way cross population in *Gossypium hirsutum* L. *Theor. Appl. Genet.* **117**, 883–894.

Richmond, T.A. and Somerville, C.R. (2000) The cellulose synthase superfamily. *Plant Physiol.* **124**, 495–498.

Ruan, Y.L., Chourey, P.S., Delmer, D.P. and Perez-Grau, L. (1997) The differential expression of sucrose synthase in relation to diverse patterns of carbon partitioning in developing cotton seed. *Plant Physiol.* **115**, 375–385.

Ruan, Y.L., Llewellyn, D.J. and Furbank, R.T. (2000) Pathway and control of sucrose import into initiating cotton fibers. *Aust. J. Plant Physiol.* **27**, 795–800.

Ruan, Y.L., Llewellyn, D. and Furbank, R.T. (2001) The control of single-celled cotton fiber elongation by developmentally reversible gating of plasmodesmata and coordinated expression of sucrose and K+ transporters and expansion. *Plant Cell* **13**, 47–60.

Ruan, Y.L., Llewellyn, D.J. and Furbank, R.T. (2003) Suppression of sucrose synthase gene expression represses cotton fiber cell Initiation, elongation, and seed development. *Plant Cell* **15**, 952–964.

Ruan, Y.L., Llewellyn, D.J., Furbank, R.T. and Chourey, P.S. (2005) The delayed initiation and shortened elongation of fuzz-like short fibers in relation to altered patterns of Sucrose synthase expression and plasmodesmata gating in a lintless mutant. *J. Exp. Bot.* **56**, 977–984.

Ruan, Y.L., Furbank, R., Brill, E., Van Thournout, M. and Arioli, A. (2008) Identification of a novel type of sucrose synthase and use thereof in fiber modification. Patent: WO 2008012058-A 31-JAN-2008. Bayer BioScience N.V. (BE); Commonwealth Scientific and Industrial Research Organization (AU).

Ruan, Y.L., Jin, Y., Li, G.J., Yang, Y.J. and Boyer, J.S. (2010) Sugar input, metabolism and signaling mediated by invertase: roles in development, yield potential and response to drought and heat. *Mol. Plant*, **3**, 942–955.

Shen, X.L., Guo, W.Z., Yuan, Y.L., Yu, J.Z., Kohel, R.J. and Zhang, T.Z. (2005) Molecular mapping of QTLs for fiber qualities in three diverse lines in Upland cotton using SSR markers. *Mol. Breed.* **15**, 169–181.

Shen, X.L., Guo, W.Z., Lu, Q.X., Zhu, X.F. and Yuan, Y.L. (2007) Genetic mapping of quantitative trait loci for fiber quality and yield trait by RIL approach in Upland cotton. *Euphytica*, **155**, 371–380.

Tang, G.Q. and Sturm, A. (1999) Antisense repression of sucrose synthase in carrot (*Daucus carota* L.) affects growth rather than sucrose partitioning. *Plant Mol. Biol.* **41**, 465–479.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and Higgins, D.G. (1997) The ClustalX windows interface: xexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucl. Acids Res.* **24**, 4876–4882.

Van Ooijen, J.W. and Voorrips, R.E. (2001) *JoinMapR Version 3.0: Software for the Calculation of Genetic Linkage Maps*. Wageningen: CPRO-DLO.

Waterhouse, P.M., Wang, M.B. and Finnegan, E.J. (2001) Role of short RNAs in gene silencing. *Trends Plant Sci.* **6**, 297–301.

Zhu, H.Y., Han, X.Y., Lv, J.H., Zhao, L., Xu, X.Y., Zhang, T.Z. and Guo, W.Z. (2011) Structure, expression differentiation and evolution of duplicated fiber developmental genes in *Gossypium barbadense* and *G. hirsutum*. *BMC Plant Biol.* **11**, 40 (doi:10.1186/1471-2229-11-40).

Zrenner, R., Salanoubat, M., Willmitzer, L. and Sonnewald, U. (1995) Evidence of the crucial role of sucrose synthase for sink strength using transgenic potato plants (*Solanum tuberosum* L.). *Plant J.* **7**, 97–107.

Supporting information

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

Figure S1 Homology of GhSusA1 to other sucrose synthases.
Figure S2 Comparison of deduced amino acid sequences of plant sucrose synthase.
Figure S3 Purification of GhSusA1 recombination protein in *E. coli* and the enzyme activity assay.
Figure S4 Genomic evolution and molecular mapping of *GhSusA1*.
Figure S5 Structures of the *GhSusA1* gene in four cotton species.
Figure S6 *Agrobacterium*-mediated transformation and plant regeneration of *G. hirsutum* W0 using hypocotyls and leaves as explants.
Figure S7 Construction and molecular identification of the sense *GhSusA1* plant expression vector.
Figure S8 Expression of *GhSus1* in *GhSusA1* transgenic lines and wild type plants.
Table S1 Primers for TAIL-PCR
Appendix S1 The information of the Q-PCR analysis based on the MIQE checklist

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.