Functional analysis of GhCHS, GhANR and GhLAR in colored fiber formation of Gossypium hirsutum L.

Gao Jianfang
Zhejiang Sci-Tech University
Shen Li
Zhejiang Sci-Tech University
Yuan Jingli
Zhejiang Sci-Tech University
Zheng Hongli
Zhejiang Sci-Tech University
Su Quansheng
Zhejiang Sci-Tech University
Yang Weiguang
Zhejiang Sci-Tech University
Zhang Liqiang
Zhejiang Sci-Tech University
Nnaemeka Ekene Vitalis
Zhejiang Sci-Tech University
Jie Sun
Shihezi University College of Agriculture
Ke Liping
Zhejiang Sci-Tech University
Yuqiang Sun (sunyuqiang@zstu.edu.cn)
Zhejiang Sci-Tech University https://orcid.org/0000-0002-9178-2487

Research article

Keywords: G. hirsutum L., Fiber color, Anthocyanidin, GhCHS, GhANR, GhLAR

Posted Date: October 4th, 2019

DOI: https://doi.org/10.21203/rs.2.11140/v2

Version of Record: A version of this preprint was published on October 29th, 2019. See the published version at https://doi.org/10.1186/s12870-019-2065-7.

Abstract

Background The formation of natural colored fibers mainly results from the accumulation of different anthocyanidins and their derivatives in the fibers of Gossypium hirsutum L. Chalcone synthase (CHS) is the first committed enzyme of flavonoid biosynthesis, and anthocyanidins are transported into fiber cell after biosynthesis mainly by Anthocyanidin reductase (ANR) and Leucoanthocyanidin reductase (LAR) to present diverse colors with distinct stability. The biochemical and molecular mechanism of pigment formation in natural colored cotton fiber is not clear. Results The three key genes of GhCHS , GhANR and GhLAR were predominantly expressed in the developing fibers of colored cotton. In the GhCHSi , GhANRi and GhLARi transgenic cottons, the expression levels of GhCHS , GhANR and GhLAR significantly decreased in the developing cotton fiber, negatively correlated with the content of anthocyanidins and the color depth of cotton fiber. In colored cotton Zongxu1 (ZX1) and the GhCHSi , GhANRi and GhLARi transgenic lines of ZX1, HZ and ZH, the anthocyanidin contents of the leaves, cotton kernels, the mixture of fiber and seedcoat were all changed and positively correlated with the fiber color. Conclusion The three genes of GhCHS , GhANR and GhLAR were predominantly expressed early in developing colored cotton fibers and identified to be a key genes of cotton fiber color formation. The expression levels of the three genes affected the anthocyanidin contents and fiber color depth. So the three genes played a crucial part in cotton fiber color formation and has important significant to improve natural colored cotton quality and create new colored cotton germplasm resources by genetic engineering.

Background

Cotton, as one of the most important economic crops, provides more than 50% of the fiber source in the textile industry [1]. However, in the textile industry, the printing and dyeing processes contain many carcinogens, resulted in bleaching difficulties, containing high concentrations of halides organic matter, most organic halides are carcinogenic, teratogenic and mutagenic [2-6] are also typical persistent organic pollutants [7,8] , which are very difficult to recover. Fortunately, the natural colored cotton fibers yarn without or very less dyeing directly into cloth, matching the increasing great demand for green products, environmental protection and human health in modern society, green products with natural colors have attracted interest in terms of their potential use in the textile industry, so the natural colored cotton and it's fabrics for it's green, ecological and eco-friendly characteristics are praised as "21st century Darling", is also undoubtedly becoming an important choice and way for the transformation and upgrading of the textile industry in China. But currently only two types of colored cotton with brown and green color are available in the actual production and textile industry, which seriously restricts the development of colored cotton industry [8-15]. Natural colored cotton undergoes pigmentation by synthesizing and accumulating natural pigments in developing fiber, biochemical analyses suggested that flavonoids were involved in the brown coloration of cotton fibers [3,16], proanthocaynidins biosynthesis and accumulation were responsible for the brown coloration in cotton fibers [10,11,17-21]. Because of the complex pigment composition and structure in colored cotton fibers, it is difficult to genetically improve cotton fiber color, and even more the mechanism of pigment formation in the colored fiber is still unclear.

Flavonoids are one of the largest groups of secondary metabolites and widely distributed in plants ranging from spermatophytes to mosses [22-24]. Anthocyanins are an important class of flavonoids that represent a large group of plant secondary metabolites. Anthocyanins are glycosylated polyphenolic compounds with a range of colors varying from orange, red, and purple to blue in flowers, seeds, fruits and vegetative tissues [25]. The main classes of these phenylpropanoid pathway derivatives include flavonols, anthocyanins, and proanthocyanidins (PAs). As water-soluble, natural pigments, anthocyanins are responsible for the red, purple and blue colors of many flowers and fruit that attract pollinators and seed dispersers [26]. Plant flowers and fruits have a variety of colors because they are closely related to anthocyanins. Over 600 anthocyanins have been identified in nature [27]. In plants, the most common anthocyanins are the derivatives of six widespread anthocyanidins, namely pelargonidin, cyanidin, delphinidin, peonidin, petunidin, and malvidin [28], which constitute the core anthocyanidins predominant in higher plants [29]. The proanthocyanins play an important role in regulating many biological stresses and abiotic stresses in plant, and play a crucial part in the physiological processes such as anti-ultraviolet, disease resistance, scavenging free radicals, regulating seed dormancy and germination [29-34]. The anthocyanin biosynthetic pathway is an extension of the general flavonoid pathway, which starts with the chalcone synthase (CHS) mediated synthesis of naringenin chalcone from 4-coumaroyl-CoA and malonyl-CoA. CHS is the key enzyme that encodes the first step of anthocyanin synthesis. After biosynthesis, flavonoids are transported to vacuoles or cell walls [35]. The last steps of monomer biosynthesis are catalyzed by two distinct enzymes. For the biosynthesis of 2, 3-trans-(6)-flavan-3-ols (e.g. catechin), leucoanthocyanidins are reduced directly to the corresponding flavan-3-ol [36,37] by leucoanthocyanidin reductase (LAR), and thus is the first committed step in PA biosynthesis. For the biosynthesis of the 2, 3-cis-type compounds (e.g. epicatechin), leucoanthocyanidins are converted to anthocyanidins by anthocyanidin synthase (ANS) and then reduced by anthocyanidin reductase (ANR) to make the corresponding 2, 3-cis-flavan3-ol [38]. The anthocyanin branchpoint enzyme UDP-glycose: flavonoid-3-O-glycosyltransferase (UF3GT) and the PA branchpoint enzyme anthocyanidin reductase (ANR) both utilize the unstable flavonoid precursor cyanidin as a substrate. Recently, LAR also was shown to convert 4b-(S-cysteinyl)-epicatechin to free epicatechin in *Medicago truncatula* and plays an important role in regulating the length of PA polymers [39]. Both LAR and ANR are NADPH/NADH-dependent isoflavone-like reductases belonging to the reductase epimerase-dehydrogenase superfamily.

So engineering paler color has been achieved relatively easily by silencing structural genes in the anthocyanin biosynthetic pathway. Shifts in color from blue to red have been achieved by silencing Flavanone 3, 5-Hydroxylase gene (*F3'5'H*) [40]. Novel, red colored seeds of soybean have been produced by inhibiting the activity of anthocyanidin reductase (ANR) in the seed coat [41]. High-level suppression of soybean *ANR1* and *ANR2* genes confers a red-brown grain phenotype, and redirects metabolic flux from PA biosynthesis into the anthocyanin and flavonol-3-*O*-glucoside pathways. ANR removes anthocyanidins to supply epicatechin for proanthocyanidin synthesis. In the absence of ANR activity, red cyanidin-based anthocyanins accumulate in the seed coat. LAR and ANR are the key enzymes of anthocyanin transport and proanthocyanidin synthesis.

The suppression of PA branchpoint genes in soybean seed coat tissue as a novel approach for engineering pigmentation in plants. As the seed coats of other economically significant crop plants also accumulate PAs (e.g. canola (*Brassica napus* L.), flax (*Linum usitatissimum*), and wheat (*Triticum* spp.) [42-44], it remains a possibility that *ANR* gene suppression could be used to color the genetically modified grains of other important crop species.

In this study, the three key genes for anthocyanidin biosynthesis and transport in natural colored cotton fiber were analyzed. The fiber color was altered with the decreased transcript level of three key genes, which resulted in the content of anthocyanidins change. It is very important to genetic manipulation *GhCHS*, *GhANR* and *GhLAR* in the anthocyanin metabolic pathway to improve the cotton fiber color, to satisfy the increasing great demand for green textile industry.

Results

Identification and phylogenetic analysis of *GhCHS*, *GhLAR* and *GhANR* genes

The differentially expressed genes were scanned from the transcriptome of brown cotton and its near isogene line [11]. From the differentially expressed genes, the genes in the anthocyanidin biosynthesis pathway including *GhCHS*, *GhLAR* and *GhANR* were selected for further analysis in the colored fiber in *G. hirsutum*. In *G. hirsutum*, 7 *GhCHS* genes and 6 *GhCHS-like* genes were scanned, 2 *GhANR* genes and 3 *GhLAR* genes were obtained (Fig. 2).

Multiple ChCHSs contained high amino acid homology, the homology of special motifs reached 100%, the *GhCHS* genes kept highly conserved in *G. hirsutum* (Fig. 2A). The *GhLAR* genes and *GhANR* genes were also much conserved in *G. hirsutum* (Fig. 2B, C). The members in the GhCHS family except GhCHSL-2 had two domains and were mostly divided into N-terminus and C-terminus, but GhCHSL-2 has only N-terminal one (Fig. 3A). The GhANR1 and GhANR2 were also divided into N-terminal and C-terminal (Fig. 3B). GhLARs had only one N-terminal domain (Fig. 3C).

Expression pattern of *GhCHS*, *GhLAR* and *GhANR* in the developing fibers

The 7 *GhCHS* genes and 6 *GhCHS-like* genes, 2 *GhANR* genes and 3 *GhLAR* transcript levels in the developing fibers of different stages in the natural colored cotton Zongxu1 (ZX1) and different cotton species were measured. The 3 *GhCHS* genes (named *GhCHS1*, *GhCHS2*, *GhCHS3*) were detected in the developing fibers of ZX1, and *GhCHS2* were predominant. The expression level of *GhCHS2* was extremely higher than *GhCHS1* and *GhCHS3* and appeared most especially in the fiber of 5 and 10 DPA (days post anthesis) (Fig. 4A). The two *GhANR* genes (*GhANR1* and *GhANR2*) were quantified in the developing fibers of ZX1, the maximal expression level appeared in the fiber of 10 DPA (Fig. 4B). The expression levels of *GhANR* genes were extremely higher than those of *GhLAR* genes to about 10-fold in the fibers of 5 DPA and 30-fold in the fibers of 10 DPA. All *GhLAR* genes were detected in the developing fibers from 0 DPA to 20 DPA, predominantly expressed in the developing fibers of 5 DPA and 10 DPA (Fig. 4C). From the expression pattern of *GhCHS*, *GhLAR* and *GhANR*, the 3 genes were all predominantly expressed in the

fibers of 10 DPA, the gene expression patterns were further detected in different cotton species with different colored fibers at 10 DPA.

The 5 cotton species (with white fiber or brown fiber) were used to measure *GhCHS*, *GhLAR* and *GhANR* expression levels in the developing fibers of 10 DPA. In plant, *chalcone synthase* (*CHS*) gene is the first committed step of flavonoid biosynthesis, the expression levels of *GhCHS* genes were significantly higher in the colored fibers than in the white fibers, especially in ZX1 and HZ fibers of 10 DPA (Fig. 5A). In the *GhCHS* genes*, GhCHS2* kept preferential expression and maintained the trend during the colored fiber development, *GhCHS1* was weakly expressed and the expression level of *GhCHS3* was almost negligible, so the *GhCHS2* was measured to represent the transcript levels of *GhCHS* genes in the developing fibers, and the conserved sequence was used to interfere their transcripts. The expression levels *GhANR* genes in colored cotton HZ lines with dark brown fiber were the highest among the 5 cotton species. The expression of *GhANR2* was greatly increased in the dark brown fibers of HZ compared with that in C312, HS2, ZH and ZX1; the transcript level of *GhANR1* in ZH lines with light brown was significantly higher than that in C312, HS2, and ZX1 (Fig. 5B). For anthocyanidins transport, the ANR represents the main way for anthocyanidins flowing into fiber cell in natural colored cotton, the expression level of *GhANR* genes in the developing fiber of the 5 species was extremely higher than that of *GhLAR* genes and *GhCHS* genes (Fig. 5). Compared with white cotton fibers, the transcription level of *GhANR* genes in brown cotton fibers was significantly higher than in white fibers. *GhLAR1* gene had the highest expression levels in the deep brown fibers of HZ lines among the 5 cotton species (Fig. 5C), significantly higher than the natural colored cotton ZX1, ZH and white fiber cotton C312 and HS2. So the conserved sequences of *GhANR1* and *GhANR2*, *GhLAR1* were used to interfere their transcripts.

Phenotypic analysis of transgenic RNAi colored cotton

The natural colored cotton ZX1 was used to silence the endogenous *GhCHS2*, *GhLAR1* and *GhANR* genes through CLCrV-based virus-induced gene silencing system. The positive control of transgenic *GhPDS*-RNAi plant appeared light bleaching symptoms in the leaves, stalks, cotton bolls and cotton fiber, which continued to be expressed in the whole life of cotton. The negative control of transgenic vector-free plants (ZX1 NCK) compared with the wild type only showed the shrinkage of the leaves (Fig. 6). The color of fibers in the *GhCHSi*, *GhANRi* and *GhLARi* were obviously fading in the depth of brown color (Fig. 6, 7). The fiber color in *GhANRi* plants was distinctly faded with brown color and significantly lighted compared to ZX1 (N CK), the fiber color in *GhLARi* plants became lighter in the depth of brown color, the cotton fiber color in *GhCHSi* plants was not significantly different from ZX1, *GhPDSi* and its NCK (Fig. 7A,B,C). Among the 5 cotton species, the fiber color of HZ was deeper than the other 4 cotton species, the fiber color of *GhANRi* HZ plants was obviously lighter than that of WT (HZ) and ZX1 (Fig. 7D). The fiber color in *GhANRi* ZH plants also obviously became lighter than WT (ZH) and ZX1 (Fig. 7E). It indicated that *GhANR* and *GhLAR* played an important role in the anthocyanin synthesis and the accumulation of pigment in cotton fiber.

Expression analysis of *GhCHS*, *GhANR* and *GhLAR* in RNAi plants

In the gene-silenced ZX1 plants, the expression levels of *GhANR*, *GhLAR* and *GhCHS* were significantly changed compared to the WT ZX1 and *GhPDSi* ZX1 plants (Fig. 8). In the *GhCHSi* ZX1 plants, the expression level of *GhCHS2* in the fibers at 5 DPA and 15 DPA was significantly lower than that of WT ZX1 and *GhPDSi* ZX1, especially in the developing fiber of 5 DPA (Fig. 8A). The expression level of *GhLAR* in the *GhCHSi* ZX1 plants appeared no significant change (Fig. S1A; see Additional file 2); the expression level of *GhANR* in the developing fiber of 15 DPA were significantly decreased (Fig. S1B; see Additional file 2). The brown color of fiber in *GhCHSi* ZX1 plants was lightly fading (Fig. 6, 7). In the *GhANRi* ZX1 plants, the *GhANR* expression level in the developing fibers of 5 DPA, 10 DPA and15DPA were significantly lower than WT ZX1 and *GhPDSi* ZX1 (Fig. 8B), the expression level of *GhCHS* in the fibers of 15 DPA increased, but had no significant change in the fibers of 5 DPA and 10DPA (Fig. S1C; see Additional file 2). The expression level of *GhLAR* had various changes (Fig. S1D; see Additional file 2). The color of brown fiber in *GhANRi* ZX1 plants was strongly fading (Fig. 6, 7). Compared with WT ZX1 and *GhPDSi* ZX1 plants, the expression level of *GhLAR* in the fiber of 5 DPA in the *GhLARi* plants was not all significantly changed, but was markedly increased in the fiber at 10 DPA, sharply decreased in the fibers at 15 DPA (Fig. 8C), the expression level of *GhCHS* in the fibers of 15 DPA was significantly increased (Fig. S1E; see Additional file 2), the expression level of *GhANR* in the fiber of 15 DPA was significantly increased in *GhLARi*-1 plant (Fig. S1F; see Additional file 2). The color of brown fiber in *GhLARi* ZX1 plants was significantly fading (Fig. 6, 7). From the *GhANRi* and *GhLARi* ZX1 plants, the suppression of *GhANR* and *GhLAR* could upregulate the expression of *GhCHS* gene (Fig. S1; see Additional file 2).

The anthocyanin content in plant tissues positively correlated with fiber color

Natural colored cotton ZX1 was large-area planted with brown fiber, here was used to study the effect of *GhANR*, *GhLAR* and *GhCHS* expression on anthocyanidins accumulation. The content of anthocyanidins of cotton kernel, fiber and seedcoat in WT ZX1 was significantly higher than those in the *GhANRi*, *GhLARi* and *GhCHSi* ZX1 plants (Fig. 9B, C). The contents of anthocyanidins in the RNAi plants were markedly decreased in the cotton kernels, fiber and seedcoat, but the anthocyanidins content was significantly increased in the leaves compared to those in WT ZX1 (Fig. 9A). The contents of anthocyanidins in leaves and cotton kernels of control plants with free-armed vector (N CK) were significantly higher than those in WT (Fig. 9A,C). The content of anthocyanidins in cotton kernels, fiber and seedcoat influenced the fiber color, the fiber color became fading with the anthocyanidins contents reduced in the *GhANRi*, *GhLARi* and *GhCHSi* plants, while the content of anthocyanidins in leaves were increased (Fig. 9B,C).

Discussion

Identification and expression pattern of *GhCHS*, *GhANR* and *GhLAR*

In the genome of *G. hirsutum*, the 13 *GhCHS* and *GhCHS-like* genes in the *CHS* family, 2 *GhANR* genes and 3 *GhLAR* genes were scanned, the gene and protein sequences of *GhCHS*, *GhANR* and *GhLAR* were highly conserved, but the genes of *GhCHS*, *GhANR* and *GhLAR* had the expression specificity in cotton plant, *GhCHS2* gene was predominantly expressed in colored cotton fibers, *GhCHS1* and *GhCHS3* expressed

weakly in the developing fibers, the other *GhCHS* and *GhCHS-like* transcripts in the developing fibers were not measured. *GhLAR1*, *GhLAR2* and *GhLAR3* were all expressed in the developing fibers, but differentially expressed in the different cotton species with different colors or color-depth, the 3 *GhLAR* genes represented the high expressive abundance in the deeply colored fibers of HZ, and perhaps the *GhLAR* genes could improve the fiber color depth. The 2 *GhANR* genes were expressed in the developing fibers and obviously increased their transcripts in the colored cotton species, and also showed high expression abundance in the deeply colored fibers of HZ. Among the three types of genes for anthocyanidin biosynthesis and transport, the *GhANR* genes always maintained high expression level and as well represented the main flow way for anthocyanidins into fiber cell [11] and played the major role for anthocyanidins transport.

The expression levels of *GhCHS*, *GhANR* and *GhLAR* closely related to fiber color

The 5 cotton species were used to measure the influence of *GhCHS*, *GhANR* and *GhLAR* gene expression on the fiber color formation. The expression levels of *GhCHSs* and *GhANRs*, *GhLAR1* and *GhLAR3* were all predominantly expressed early in developing fibers of colored fibers, especially in the dark brown fiber of HZ (Fig. 5). The expression levels of *GhCHS*, *GhANR* and *GhLAR* positively influenced the color formation of fiber in colored cotton. Therefore, for improving the color of cotton fiber, firstly the *GhCHS* gene expression would be increased to enhance the anthocyanin biosynthesis, then the *GhANR* and *GhLAR* increased their expression for transporting anthocyanidins into fiber cell. In the *GhANRi* and *GhLARi* cotton lines, the *GhCHS* gene was upregulated by the suppression of *GhANRi* and *GhLARi*, perhaps in natural colored cotton, the PA formation in the fiber cell could feedback the anthocyanidins biosynthesis, PA formation in fiber cell was mainly resulted from the anthocyanidin transport and accumulation through GhANR and GhLAR. Correspondingly, the suppression of *GhCHS* in *GhCHSi* cotton lines, the *GhANR* was downregulated, perhaps no more anthocyanidins could be transported into fiber cell through GhANR. The content of anthocaynidins in cotton kernels, fiber and seedcoat of *GhANRi*, *GhLARi* and *GhCHSi* plants decreased and increased pattern in leaves could confirm this hypothesis.

The suppression of *GhCHS*, *GhANR* and *GhLAR* had negative effect on fiber color

The *GhANR*, *GhLAR* and *GhCHS* genes in natural colored cotton ZX1 was silenced, and the fiber color in the transgenic RNAi ZX1 plants was significantly different from the WT and CK. In the transgenic ZX1 plants, the endogenous genes of *GhANR*, *GhLAR* and *GhCHS* were suppressed, especially in the fiber of 5 DPA and 10 DPA (Fig 8), the fiber color in the transgenic ZX1 plants faded to lighter and even more lighter. The down-regulation levels of the 3 genes emerged as negative correlation with fiber color. In the general phenylpropanoid pathway, chalcone synthase was the first committed enzyme of flavonoid biosynthesis, among the 3 genes, the conserved sequence of *GhCHS1, GhCHS2* and *GhCHS3* silenced has little significant effect on cotton fiber color. Firstly, it may be multiple members of *CHS* family in *G. hirsutum*, although *GhCHS2* predominantly expressed early in developing fiber in colored cotton, other members existed functional complementarity after *GhCHS2*, even *GhCHS1* and *GhCHS3* suppressed; Secondly, *GhCHS* genes were in the upstream location of anthocyanidin biosynthesis, suppression of *GhCHS* had

little effect on downstream synthesis and metabolism of anthocyanins. Early biosynthetic genes (EBGs)—*CHS*, *CHI*, and *F3H* are the common flavonoid pathway genes which are involved in the biosynthesis of all downstream flavonoids. In general, the reported expression profile of EBGs varies and there is no consistent correlation between their expression levels and anthocyanin content in *Solanaceous* vegetables [51]. In eggplant, the expression level of *SmCHS* was reported to be significantly upregulated in black or violet fruits compared to the green or white mutants [52, 53]. In potato tubers, the association of expression of *CHS* and anthocyanin accumulation is more consistent. *CHS* were highly expressed in red and purple tubers and correlated with anthocyanin content [54-57].

The *GhANR* and *GhLAR* worked for anthocyanidins transport in the anthocyanin metabolic pathway, the *GhANR* played the main role for colored anthocyanidins into fiber cell, the *GhLAR* worked for transporting leucoanthocyanidin in fiber cell and also could enhance the fiber color perhaps by polymerization and oxidation to form anthocyanin derivatives [11]. The *GhLARs* were preferentially expressed in the deep colored fiber of HZ plant, the fiber color became lighter in the *GhLAR* suppressed plants.

PAs (also called condensed tannins) are synthesized via a branch of anthocyanin biosynthesis pathway under the catalyzation of leucoanthocyanidin reductase (LAR) and anthocyanidin reductase (ANR). LAR catalyzes the conversion of leucoanthocyanidin (flavan-3, 4-diol) to catechin, while ANR catalyzes the synthesis of epicatechin from anthocyanidin [36, 38, 58]. Ectopic expression of the tea *CsLAR* gene in tobacco results in the accumulation of higher level of epicatechin than that of catechin, suggesting LAR maybe involved in the biosynthesis of epicatechin [37]. ANRs from grapevine and tea are proven to have epimerase activity and thus can convert anthocyanidin to a mixture of epicatechin and catechin [37, 59]. Further, previous engineering experiments in soybean, *Arabidopsis*, and *petunia* have redirected metabolic flux from anthocyanin biosynthesis into the isoflavone pathway, from lignin biosynthesis into the flavonoid pathway, and from flavonol biosynthesis into the anthocyanin pathway, by suppressing anthocyanin, lignin, and flavonol branchpoint genes respectively [60-62]. The overexpression of the *ANR* gene from *Medicago truncatula* in tobacco resulted in reduced anthocyanin pigmentation in the flower and elevated PA levels [58]. These results suggested the potential for *ANR* to compete with the anthocyanin biosynthesis enzyme UDP-glycose: flavonoid-3-*O*-glycosyltransferase (UF3GT) for the substrate anthocyanidin, suppression of *ANR* genes results in increased anthocyanin accumulations.

The *Arabidopsis ANR* (or *BAN*) knockout mutant displayed precocious accumulation of cyanic pigments in the seed coat during early seed development [63]. The accumulations were only temporary, and resulted in a transparent *testa* (*tt*) phenotype with black pigmentation confined to the raphe of the dried grain [63]. This contrasts the phenotype in soybean, where high-level suppression of ANR genes gives a red-brown grain [41]. There may exist underlying mechanistic and metabolite differences that could explain the differences in grain phenotypes between these species. In *Arabidopsis*, the UF3GT gene (*UGT78D2*) and the *ANR* gene are regulated reciprocally, with *UGT78D2* expressed with anthocyanins in the seedling, and *ANR* expressed with PAs in the seed coat [64]. By contrast, soybean UF3GT genes (*UGT78K1* and *UGT78K2*) and ANR genes (*ANR1* and *ANR2*) are both expressed in the seed coat [41]. Thus, it is possible that the difference in phenotype between the soybean grain undergoing high-level ANR

gene suppressions and the *Arabidopsis ANR* knockout grain, may be attributed to the presence and absence of UF3GT expressions, respectively, that stabilize anthocyanins in soybean allowing their accumulations to provide the red-brown grain phenotype.

The biosynthesis of flavan-3-ols has been well characterized in many plant species, both genetically and biochemically. Two distinct enzymes, LAR and ANR, are involved in catalyzing the last steps of the pathway to flavan-3-ol monomers in PA-producing plants [37, 65, 66]. Genes encoding LAR and ANR can occur as single gene, for example in *Arabidopsis* [38], or as multigene families, for example in grapevine [65] and tea [37]. Analysis of the *P. trichocarpa* genome revealed three loci encoding LAR proteins and two loci encoding ANR proteins [67, 68], the enzymatic activity of the proteins encoded by all loci by heterologous expression and *in vitro* enzyme assays and showed that they are likely involved in the catalysis of the last steps of flavan-3-ol biosynthesis in native black poplar. ANRs and LARs are two distinct classes of enzymes and that DFR is more related to ANRs than LARs. Similar evolutionary relationships for ANR and LAR proteins were reported [37, 68]. Transcript levels of all three *PnLAR* and two *PnANR* genes increased in rust-infected black poplar leaves over the course of infection. Monomeric catechin synthesized from the LAR branch is freely available and accumulated in black poplar, while free ANR-dependent epicatechin was observed only at very low concentrations. The recovery of epicatechin after hydrolysis of PAs indicates that epicatechin might contribute to the extension of PA chains. Similar mechanisms also were observed in grape and Norway spruce [65, 69]. LARs promoted the biosynthesis of catechin monomers and inhibited their polymerization. The accumulation of catechin monomers and polymers was increased by up-regulating the expression of *NtLAR* and *NtANR* s in *CsMYB5b* transgenic tobacco [70]. So the transport of anthocyanidins through GhANR, GhLAR into fiber cell will be the important link for genetic engineering of colored fiber molecular improvement.

The anthocyanidins content in the fiber directly influenced fiber color

In the transgenic *RNAi* cotton plants, the content of anthocyanidins was reduced by suppression of the endogenous *GhANR*, *GhLAR* and *GhCHS* genes, which resulted in the fiber color fading. CHS plays an important role in the phenylalanine metabolic pathway, plant growth and development, such as stress response, plant fertility and plant color [71]. LAR is a key enzyme in the synthetic pathway of plant flavonoids from phenylalanine, which catalyzes the conversion of colorless anthocyanins to catechins [58, 65, 66]. Transcript levels of LAR1 and ANR2 genes were significantly correlated with the contents of catechin and epicatechin to regulate PA synthesis, respectively. Ectopic expression of apple *MdLAR1* gene in tobacco suppresses expression of the late genes in anthocyanin biosynthetic pathway, resulting in loss of anthocyanin in flowers [66].

The anthocyanidins content in the fiber and seedcoat of *GhLARi* plants was higher than *GhANRi* plants, and the fiber color was also deeper than that of *GhANRi* plant, although LAR transported colorless anthocyanins into fiber cell. From our previous research, the transcription level of *GhLAR* in the fibers of brown cotton was higher than that in white cotton, during the fiber development, the fiber color of *GhLARi* plants lightly faded here. Compared with white cotton fibers, the expression level of *GhANR* in brown

cotton fibers was significantly higher. The gene expression of *GhANR* was active in brown cotton fibers and reached its peak at 12 DPA, when the expression level of *GhANR* in brown cotton fibers was >7 times higher than that in white cotton fibers [11]. During the fiber development, the *GhLAR* expression level in brown cotton was much lower than that of *GhANR*, so effect of suppression of *GhLAR* on the fiber color change was lower than that of *GhANR*, the suppression of *GhANR* in ZX1 could cause the fiber color to be significantly lighter. Our work of NMR analyses demonstrated that the flavan-3-ols in brown and white cotton fibers were in the 2, 3-cis form, but part of the proanthocyanidins in the white cotton fibers were modified by acylation. The prodelphidin (PD) relative percentage was similar to that of procyanidin (PC) in white cotton fibers, and proanthocyanidins with 90.1% PD were found in brown cotton fibers. The proanthocyanidin monomeric composition was consistent with the expression profiles of proanthocyanidin synthase genes, suggesting that ANR represented the major flow of the proanthocyanidin biosynthesis pathway in brown cotton fibers. Compared with white fibers, all of the proanthocyanidin synthase genes were expressed at a higher level in brown fibers [11]. The cis-form and trans-form of flavan-3-ols were synthesized via leucoanthocyanidin reductase (LAR) and anthocyanidin reductase (ANR) branches, respectively [11, 38, 58, 72]. Biochemical analyses by mass spectrometry (MS) revealed that the main PA monomers in brown cotton fibers contained three hydroxyls on the B ring (gallocatechin or epigallocatechin) [11,21,73], PA accumulation in brown fibers starts at an early stage (5 DPA) and peaks at 30 DPA, whereas in mature brown fibers, PAs are converted to oxidized derivatives (quinones). Because developing brown fibers do not exhibit distinct coloration until maturation, the condensed quinones were proposed instead of their PA precursors, directly contribute to brown pigmentation in cotton fibers [11]. Therefore, the three key genes in the anthocyanin metabolic pathways played the very important role in the coloration of cotton fibers, and became the target genes for genetic manipulation to improve cotton fiber color.

Conclusions

In colored cotton fibers of *G. hirsutum*, *GhCHS2* gene was predominantly expressed in developing colored cotton fibers among 7 *GhCHS* and 6 *GhCHS-like* genes and represented *CHS* gene in anthocyanin metabolism in colored fibers. 2 *GhANR* genes and 3 *GhLAR* genes were highly conserved and homologous, significantly expressed in the developing colored cotton fibers. The *GhCHS2*, *GhANR* and *GhLAR* genes were differentially expressed in the colored cotton fibers with different color depth. The *GhCHS*, *GhANR* and *GhLAR* genes were interfered in colored cottons with different color depth, the expression levels of the three genes were significantly declined, the anthocyanin contents in the RNAi cotton plants were significantly reduced with the declined gene expression, and the fiber color was significantly changed and weaken. The three genes of *GhCHS*, *GhANR* and *GhLAR* played a crucial part in cotton fiber color formation and has important significant to improve natural colored cotton quality through genetic manipulation of the three genes and create new colored cotton germplasm resources by genetic engineering.

Methods

Plant materials

The *G. hirsutum* L. cv. Coker 312 (C312) and HS2 with white fiber, natural colored cotton ZX1 (Zongxu 1) with brown fiber, HZ with dark brown fiber and ZH with lighter brown fiber were used in this study (Fig. 1). The cotton seeds *G. hirsutum* cv. C312, HS2, ZH and HZ were preserved at the Key Laboratory for Plant Secondary metabolism and regulation of Zhejiang Province, Zhejiang Sci-Tech University, Hangzhou, China. Cotton seeds of ZX1 were kindly provided by Dr. Xiongming Du (Institute of Cotton Research, CAAS). Seeds were germinated and grown in a greenhouse at 28°C with a 14 h light and 10 h dark cycle. Seedlings with a 2^{nd} true leaf emergence were used for agroinfiltration. Infiltrated plants were grown in the greenhouse at 23-25°C with a 14/10 h light/dark photoperiod. The cotton plants were cultivated in the field under standard conditions. The samples of cottonseed, fiber and seedcoat were collected at the time of 0 DPA, 5 DPA, 10 DPA, 15 DPA and 20 DPA (days post anthesis) respectively (the fiber and seedcoat at 10, 15 and 20 DPA were removed from the cotton seed kernel), then put into liquid nitrogen and stored in the -80℃ ultra-low temperature freezer for RNA extraction.

Gene cloning and construction of RNAi vectors

The candidate genes were obtained from the differentially expressed genes in brown cotton and its near isogene line (*G. hirsutum*) [11]. The *GhANR*, *GhLAR* and *GhCHS* genes were scanned in the cotton genomes (http://www.cottongen.org) for gene accesses and sequences. BioEdit was used for multiple sequence alignment with the amino acid sequences of *GhANR1*, *GhANR2*, *GhLAR1*, *GhLAR2*, *GhLAR3*, and *GhCHS* genes. The characteristics of the genes coded proteins were used TMHMM (http://www.cbs.dtu.dk/services/TMHMM-2.0/) analysis for the transmembrane region of the protein encoded by mRNA. The components, physicochemical properties and isoelectric points of amino acid sequences are analyzed by ProtParam (http://web.expasy.org/ protparam/), respectively.

The new CLCrV-based vector was modified from the CLCrV DNA-A and DNA-B components individually, which were inserted into the pCambia1300 vector to generate pCLCrVA and pCLCrVB, respectively [45]. The fragments of candidate genes were inserted into pCLCrVA to produce pCLCrVA-*GhCHS*, pCLCrVA-*GhANR,* pCLCrVA-*GhLAR* for VIGS in cotton plants as described in the previous papers [45-47].

A 400-600 bp fragment of the candidate genes of *GhANR2*, *GhLAR1* and *GhCHS2* was amplified as our previous paper [46]. The *phytoene desaturase* gene (*PDS*) causes loss of chlorophyll and carotenoids and was used as a positive marker to visualize the timing and extent of endogenous gene silencing. A 327 bp fragment of the *PDS* gene isolated from C312 by PCR to construct vector pCLCrV-*GhPDS.*

The four vectors with or without foreign genes were transformed individually into *Agrobacterium tumefaciens* strain GV3101 by electroporation using a Gene Pulser Apparatus (Bio-Rad, Hercules, CA, USA). Three combinations, pCLCrVA-empty and pCLCrVB (for a negative control, N CK), pCLCrV-*PDS* and pCLCrVB (for a positive control, *GhPDSi*), pCLCrVA-*GhCHS*, pCLCrVA-*GhANR,* pCLCrVA-*GhLAR* and pCLCrVB (for target gene silencing) were used. Plants were transformed with pCLCrV-*PDS* and the

pCLCrVA-empty vectors as controls. The primers for cloning and detection in the experiment are listed in Additional file 1.

Cotton seedlings were grown in a growth chamber at 28℃ with a 14 h light and 10 h dark cycle. Healthy 2 week old seedlings were infiltrated with different *Agrobacteria* carrying pCLCrVA or one of its derivatives and pCLCrVB. *Agrobacteria* harboring pCLCrVA or one of its derivatives was mixed with an equal volume of *Agrobacteria* harboring pCLCrVB. The mixed *Agrobacteria* solutions were infiltrated into the abaxial side of the cotyledons of the 2-week-old cotton seedlings using syringes without needles. The agroinfiltration was repeated at least three times with at least 30 plants for each vector. Total DNA was extracted from the leaves of pCLCrV-inoculated cotton plants. The presence of pCLCrV DNA in infected plants was detected by PCR using primers specific for either pCLCrV DNA-A (CLCrVA F and CLCrVA R) as the described before [46]. Plants infiltrated with the pCLCrVA-empty vector and the wild-type C312 and ZX1 or HZ or ZH were used as controls in the experiment. Plants infiltrated with the pCLCrV-*PDS* vector showed the typical photobleaching phenotype in newly developing leaves, different tissues and organs. However, the efficiency of gene silencing was judged by the intensity of photobleaching during the whole growth period. The leaves and developing bolls in transgenic *GhCHSi*, *GhANRi*, *GhLARi*, *GhPDSi* and CKs plants were respectively collected at 0 DPA, 5 DPA, 10 DPA and 15 DPA for measurement of anthocyanin content and gene expression analysis.

Gene expression analysis by quantitative real-time PCR

Total RNA was isolated from the leaves, the mixture of fiber and seedcoat according to the manufacturer's instructions (RNAprep Pure TIANGEN BIOTECH, China), and treated extensively with RNase-free DNase I. Double-stranded cDNA was synthesized from 200ng RNA using FastQuant RT kit with gDNase (TIANGEN BIOTECH, China) according to a standard double-stranded cDNA synthesis protocol. Real-time PCR (qRT-PCR) assays were performed using the SYBR FAST qPCR kit (KAPA SYBR®, USA) and the qRT-PCR reaction was performed using the ABI QS3 fluorescence quantitative PCR instrument (ABI, USA). Specificity of the amplified PCR product was determined based on melting curve analysis. Primers for target genes were designed using Primer Express 5 (Premier Biosoft, Palo Alto, CA) and are listed in Table 1. The cotton *Ubiquitin7* gene (*GhUBQ7*, Gen Bank accession number: DQ116441) was used as an internal control for the assays. The expression of *GhANR* genes, *GhLAR* genes and *GhCHS* genes in cotton were obtained and standardized to the constitutive *GhUBQ7* gene expression level.

The analysis of anthocyanin content of transgenic plants

Measurements of anthocyanidin accumulation were performed as described by Jeong et al. (2010) [48] and Wade et al. (2003) [49]. Weighed samples (approx. 100 mg) in a 1.5 mL microfuge tubes were harvested into liquid nitrogen to freeze plant tissue. Samples were extracted overnight in 1ml of 0.5% (v/v) HCl in methanol, and then violently shaken in vortex for 30 sec. The extraction buffers were shaken with 120 rpm in the dark for 1 hour. The extraction buffers were centrifuged at 2,630g for 15 min at 20℃. This process was repeated 3 times. The supernatant was assayed spectrophotometrically (UV-2600,

Shimadzu, Japan) and anthocyanidin absorbance units (A_{530}–A_{657}) per gram fresh weight were calculated. The blank should be 480ml Methanol with 0.5% (v/v) HCl and 320ml Milli-Q H_20 for a total of 800ml. A spectrophotometer was used for the absorbance measurements at 530, 620, and 650 nm. The results were determined based on the following equation: optical density (OD) = (A_{530} – A_{620}) – [0.1 × (A_{650} – A_{620})] [50].

Statistical analysis

All data are presented as mean ± SD from at least three independent experiments with three replicates each. The statistical significance of the differences was determined using the Student's t-test. Differences between treatments were considered significant when *P < 0.05, **P <0.01 and ***P <0.005 in a two-tailed analysis.

Abbreviations

CHS: Chalcone synthase; ANR: Anthocyanidin reductase; LAR: Leucoanthocyanidin reductase; PA: Proanthocyanidins; UF3GT: UDP-glycose: flavonoid-3-O-glycosyltransferase; DPA: Days post anthesis; CLCrV: geminivirus Cotton leaf crumple virus; PDS: Phytoene desaturase

Declarations

Ethics approval and consent to participate

Not applicable.

Consent to publication

Not applicable.

Availability of data and materials

All data generated or analyzed during this study are included in this published article.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the National Key Research and Development Program of China (2018YFD0100401), and General Program of National Natural Science Foundation of China (31671738). The funding agencies had no role in research design, data collection and analysis, or manuscript writing.

Author Contributions

YS conceived and designed the experiments. JG, LS, JY, HZ, QS, WY, VEN, LZ and LK performed the research and prepared the figures 1–7 and figure 9. JS participated in data analysis. YS, QL and LK wrote and VEN corrected the article. All authors reviewed and approved the manuscript.

Acknowledgments

The authors thank Dr. Xiongming Du for providing ZX1 seeds and the helpful discussions.

References

1. Hagenbucher S, Olson DM, Ruberson J, Wäckers FL, Romeis J. Resistance mechanisms against arthropod herbivores in cotton and their interactions with natural enemies. Crit Rev Plant Sci, 2013; 32: 458–482.
2. Efe L, Killi F, Mustafayev SA. An evaluation of eco-friendly natural colored cottons regarding seed cotton yield, yield components and major lint quality traits under conditions of East Mediterranean region of Turkey. Pak J Biol Sci. 2009; 12: 1346–1352.
3. Hua S, Yuan S, Shamsi IH, Zhao X, Zhang X, Liu Y, Wen G, Wang X, Zhang H. A comparison of three isolines of cotton differing in fiber color for yield, quality, and photosynthesis. Crop Sci. 2009; 49: 983–989.
4. Kimmel LB, Day MP. New life for an old fiber: Attributes and advantages of natural colored cotton. AATCC Rev. 2001; 1: 32–36.
5. Gonzalez P, Zaror C. Effect of process modifications on AOX emissions from kraft pulp bleaching, using chilean pine and eucalyptus. J Clean Prod. 2000; 8: 233–241.
6. Sharma A, Thakur VV, Shrivastava A, Jain RK, Mathur RM, Gupta R, Kuhad RC. Xylanase and laccase based enzymatic kraft pulp bleaching reduces adsorbable organic halogen (AOX) in bleach effluents: A pilot scale study. Bioresource Techno. 2014; 169: 96–102.
7. Ramjaun SN, Yuan R, Wang Z, Liu J Degradation of reactive dyes by contact glow discharge electrolysis in the presence of Cl- ions: Kinetics and AOX formation. Electrochim Acta. 2011; 58: 364–371.
8. Yan Q, Wang Y, Li Q, Zhang Z, Ding H, Zhang Y, Liu H, Luo M, Liu D, Song W, Liu H, Yao D, Ouyang X, Li Y, Li X, Pei Y, Xiao Y. Upregulation of GhTT2-3A in cotton fibers during secondary wall thickening results in brown fibers with improved quality. J Plant Biotechnol. 2018; 16: 1735–1747.
9. Shi YZ, Du XM, Liu GQ, Qiang AD, Zhou ZL, Pan ZE, Sun JL. Genetic analysis of natural colored lint and fuzz of cotton. Cotton Sci. 2002; 14: 242–248.
10. Feng H, Tian X, Liu Y, Li Y, Zhang X, Jones BJ, Sun Y, Sun J. Analysis of flavonoids and the flavonoid structural genes in brown fiber of upland cotton. PLoS ONE. 2013; 8: e58820.
11. Feng H, Li Y, Wang S, Zhang L, Liu Y, Xue F, Sun Y, Sun J. Molecular analysis of proanthocyanidins related to pigmentation in brown cotton fibre (G*ossypium hirsutum* L.). J Exp Bot. 2014; 65: 5759–5769.

12. Tan YL, Zhou XZ. The research status and development trends of natural colored cotton. Progress. Text Sci Tech. 2015; 2: 1–4.
13. Ma M, Hussain M, Memon H, Zhou W. Structure of pigment compositions and radical scavenging activity of natural green-colored cotton fiber. Cellulose. 2016; 23: 955–963.
14. Bi YW. Performance analysis and development trend of natural colored cotton. Shandong Text Sci Tech. 2017; 4: 4–7.
15. Du XM, Liu F, Wang KB, Jia YH, Zhou ZL, He SP, Gong WF, Pan ZE, Wang LR, Geng XL, Pang BY. Collection, evaluation and utilization of cotton germplasm. Cotton Sci. 2017; 29: 51–61.
16. Murthy MS. Never say dye: the story of colored cotton. Resonance. 2001; 6: 29–35.
17. Gong W, He S, Tian J, Sun J, Pan Z, Jia Y, Sun G, Du X. Comparison of the transcriptome between two cotton lines of different fiber color and quality. PLoS ONE. 2014; 9: e112966.
18. Hinchliffe DJ, Condon BD, Thyssen G, Naoumkina M, Madison CA, Reynolds M, Delhom CD, Fang DD, Li P, McCarty J. The GhTT2_A07 gene is linked to the brown color and natural flame retardancy phenotypes of Lc1 cotton (*Gossypium hirsutum* L.) fibres. J Exp Bot. 2016; 67: 5461–5471.
19. Li YJ, Zhang XY, Wang FX, Yang CL, Liu F, Xia GX, Sun J. A comparative proteomic analysis provides insights into pigment biosynthesis in brown color fiber. J Proteomics. 2013; 78: 374–388.
20. Xiao YH, Zhang ZS, Yin MH, Luo M, Li XB, Hou L, Pei Y. Cotton flavonoid structural genes related to the pigmentation in brown fibers. Biochem Biophys Res Commun. 2007; 358: 73–78.
21. Xiao YH, Yan Q, Ding H, Luo M, Hou L, Zhang M, Yao D, Liu HS, Li X, Zhao J, Pei Y. Transcriptome and biochemical analyses revealed a detailed proanthocyanidin biosynthesis pathway in brown cotton fiber. PLoS ONE. 2014; 9: e86344.
22. Tohge T, Watanabe M, Hoefgen R, Fernie AR. The evolution of phenylpropanoid metabolism in the green lineage. Crit Rev Biochem Mol. 2013; 48: 123–152.
23. Buer C, Imin N, Djordjevic M. Flavonoids new roles for old molecules. J Int Plant Bio. 2010; 52: 98–111.
24. Cheynier V, Comte G, Davies KM, Lattanzio V, Martens S. Plant phenolics: Recent advances on their biosynthesis, genetics, and ecophysiology. Plant Physiol Bioch. 2013; 72: 1–20.
25. Tanaka Y, Ohmiya A. Seeing is believing: engineering anthocyanin and carotenoid biosynthetic pathways. Curr Opin Biotechnol. 2008; 19: 190–197.
26. Shang Y, Venail J, Mackay S, Bailey PC, Schwinn KE, Jameson PE, Martin CR, Davies KM. The molecular basis for venation patterning of pigmentation and its effect on pollinator attraction in flowers of antirrhinum. New Phytol. 2011; 189: 602–615.
27. Smeriglio A, Barreca D, Bellocco E, Trombetta D. Chemistry, pharmacology and health benefits of anthocyanins. Phytother Res. 2016; 30(8): 1265–1286.
28. Kong JM, Chia LS, Goh NK, Chia TF, Brouillard R. Analysis and biological activities of anthocyanins. Phytochem. 2003; 64: 923–933.

29. Zhang Y, Butelli E, Martin C. Engineering anthocyanin biosynthesis in plants. Curr Opin Plant Bio. 2014; 19: 81–90.
30. Feild TS, Lee DW, Holbrook NM. Why leaves turn red in autumn. The role of anthocyanins in senescing leaves of red-osier dogwood. Plant Physiol. 2001; 127: 566–574.
31. Gould KS, Mckelvie J, Markham KR. Do anthocyanins function as antioxidants in leaves? Imaging of H2O2 in red and green leaves after mechanical injury. Plant Cell Environ. 2010; 25: 1261–1269.
32. Li J, Ou-Lee TM, Raba R, Amundson RG, Last RL. Arabidopsis flavonoid mutants are hypersensitive to UV-B irradiation. Plant Cell. 1993; 5: 171–179.
33. Olsen KM, Slimestad R, Lea US, Brede C, Løvdal T, Ruoff P, Verheul M, Lillo C. Temperature and nitrogen effects on regulators and products of the flavonoid pathway: experimental and kinetic model studies. Plant cell environ. 2010; 32: 286–299.
34. Rubin G, Tohge T, Matsuda F, Saito K, Scheible W. Members of the LBD family of transcription factors repress anthocyanin synthesis and affect additional nitrogen responses in *Arabidopsis*. Plant Cell. 2009; 21: 3567–3584.
35. Koes R, Verweij W, Quattrocchio F. Flavonoids: a colorful model for the regulation and evolution of biochemical pathways. Trends Plant Sci. 200; 10: 236–242.
36. Tanner GJ, Francki KT, Abrahams S, Watson JM, Larkin PJ, Ashton AR. Proanthocyanidin biosynthesis in plants. Purification of legume leucoanthocyanidin reductase and molecular cloning of its cDNA. J Bio Chem. 2003; 278: 31647–31656.
37. Pang Y, Abeysinghe IS, He J, He XZ, Huhman D, Mewan KM, Sumner LW, Yun J, Dixon RA. Functional characterization of proanthocyanidin pathway enzymes from tea and their application for metabolic engineering. Plant Physiol. 2013; 161: 1103–1116.
38. Xie DY, Sharma SB, Dixon RA. Anthocyanidin reductases from *Medicago truncatula* and *Arabidopsis thaliana*. Arch Biochem Biophys. 2004; 422: 91–102.
39. Liu C, Wang X, Shulaev V, Dixon RA. A role for leucoanthocyanidin reductase in the extension of proanthocyanidins. Nat Plants. 2016; 2: 16182.
40. Nishihara M, Nakatsuka T. Genetic engineering of flavonoid pigments to modify flower color in floricultural plants. Biotechnol Lett. 2011; 33: 433–441.
41. Kovinich N, Saleem A, Rintoul TL, Brown DC, Arnason JT, Miki B. Coloring genetically modified soybean grains with anthocyanins by suppression of the proanthocyanidin genes ANR1 and ANR2. Transgenic Res. 2012; 21: 757–771.
42. McCallum JA, Walker JR. Proanthocyanidins in wheat bran. Cereal Chem. 1990; 67: 282–285.
43. Naczk M, Nichols T, Pink D, Sosulski F. Condensed tannins in Canola Hulls. J Agric Food Chem. 1994; 42: 2196–2200.
44. Venglat P, Xiang D, Qiu S, Stone SL, Tibiche C, Cram D, Alting-Mees M, Nowak J, Cloutier S, Deyholos M, Bekkaoui F, Sharpe A, Wang E, Rowland G, Selvaraj G, Datla R. Gene expression analysis of flax seed development. BMC Plant Biol. 2011; 11: 1–15.

45. Gu Z Huang C, Li F, Zhou X. A versatile system for functional analysis of genes and micrornas in cotton. Plant Biotech J. 2014; 12: 638–649.
46. Fu WF, Shen Y, Hao J, Wu JY, Ke LP, Wu CY, Huang K, Luo BL, Xu MF, Cheng XF, Zhou XP, Sun J, Xing CZ, Sun YQ. Acyl-CoA N-acyltransferase influences fertility by regulating lipid metabolism and jasmonic acid biogenesis in cotton. Sci Rep. 2015; 5: 11790.
47. Lu R, Martin-Hernandez A M, Peart JR, Malcuit I, Baulcombe DC. Virus-induced gene silencing in plants. Methods. 2004; 30: 296–303.
48. Jeong SW, Das PK, Jeoung SC, Song JY, Lee HK, Kim YK, Kim WJ, Park YI, Yoo SD, Choi SB, Choi G, Park YI. Ethylene suppression of sugar-induced anthocyanin pigmentation in *Arabidopsis*. Plant Physiology. 2010; 154 (3) 1514-1531.
49. Wade HK, Sohal AK, Jenkins GI. *Arabidopsis* ICX1 is a negative regulator of several pathways regulating flavonoid biosynthesis genes. Plant Physiology. 2003; 131 (2) 707-715.
50. An JP, Qu FJ, Yao JF, Wang XN, You CX, Wang XF, Hao YJ. The bZIP transcription factor MdHY5 regulates anthocyanin accumulation and nitrate assimilation in apple. Horticulture Research. 2017; 4: 17023.
51. Liu Y, Tikunov Y, Schouten RE, Marcelis L, Visser R, Bovy A. Anthocyanin biosynthesis and degradation mechanisms in *solanaceous* vegetables: a review. Front Chem. 2018; 6: 52.
52. Stommel JR, Dumm JM. Coordinated regulation of biosynthetic and regulatory genes coincides with anthocyanin accumulation in developing eggplant fruit. J Am Soc Hortic Sci. 2015; 140: 129–135.
53. Gisbert C, Dumm JM, Prohens J, Vilanova S, Stommel JR. A spontaneous eggplant (*Solanum melongena* L.) color mutant conditions anthocyanin-free fruit pigmentation. Hortsci. 2016; 51: 793–798.
54. André CM, Schafleitner R, Legay S, Lefèvre I, Aliaga CA, Nomberto G, Hoffmann L, Hausman JF, Larondelle Y, Evers D. Gene expression changes related to the production of phenolic compounds in potato tubers grown under drought stress. Phytochem. 2009; 70: 1107–1116.
55. Jung CS, Griffiths HM, De Jong DM, Cheng S, Bodis M, Kim TS, De Jong WS. The potato developer (D) locus encodes an R2R3 MYB transcription factor that regulates expression of multiple anthocyanin structural genes in tuber skin. Theor Appl Gene. 2009; 120: 45–57.
56. Payyavula RS, Singh RK, Navarre DA. Transcription factors, sucrose, and sucrose metabolic genes interact to regulate potato phenylpropanoid metabolism. J Exp Bot. 2013; 64: 5115–5131.
57. Liu Y, Lin-Wang K, Deng C, Warran B, Wang L, Yu B, Yang H, Wang J, Espley RV, Zhang J, Wang D, Allan AC. Comparative transcriptome analysis of white and purple potato to identify genes involved in anthocyanin biosynthesis. PLoS ONE. 2015; 10: e0129148.
58. Xie DY, Sharma SB, Paiva NL, Ferreira D, Dixon RA. Role of anthocyanidin reductase, encoded by BANYULS in plant flavonoid biosynthesis. Science. 2003; 299: 396–399.
59. Gargouri M, Chaudière J, Manigand C, Maugé C, Bathany K, Schmitter JM, Gallois B. The epimerase activity of anthocyanidin reductase from Vitis vinifera and its regiospecific hydride transfers. Biol Chem. 2010; 391: 219–227.

60. Yu O, Shi J, Hession AO, Maxwell CA, McGonigle B, Odell JT. Metabolic engineering to increase isoflavone biosynthesis in soybean seed. Phytochem. 2003; 63: 753–763.
61. Besseau S, Hoffmann L, Geoffroy P, Lapierre C, Pollet B, Legrand M. Flavonoid accumulation in Arabidopsis repressed in lignin synthesis affects auxin transport and plant growth. Plant Cell. 2007; 19: 148–162.
62. Nakatsuka T, Abe Y, Kakizaki Y, Yamamura S, Nishihara M. Production of red-flowered plants by genetic engineering of multiple flavonoid biosynthetic genes. Plant Cell Rep. 2007; 26: 1951–1959.
63. Albert S, Delseny M, Devic M. BANYULS, a novel negative regulator of flavonoid biosynthesis in the *Arabidopsis* seed coat. Plant J. 1997; 11:289–299.
64. Lee Y, Yoon H, Paik YS, Liu JR, Chung WI, Choi G. Reciprocal regulation of *Arabidopsis* UGT78D2 and BANYULS is critical for regulation of the metabolic flux of anthocyanidins to condensed tannins in developing seed coats. J Plant Biol. 2005; 48: 356–370.
65. Bogs J, Downey MO, Harvey JS, Ashton AR, Tanner GJ, Robinson SP. Proanthocyanidin synthesis and expression of genes encoding leucoanthocyanidin reductase and anthocyanidin reductase in developing grape berries and grapevine leaves. Plant Physiol. 2005; 139: 652–663.
66. Liao L, Vimolmangkang S, Wei G, Zhou H, Korban SS, Han Y. Molecular characterization of genes encoding leucoanthocyanidin reductase involved in proanthocyanidin biosynthesis in apple. Front Plant Sci. 2015; 6: 243.
67. Yuan L, Wang L, Han Z, Jiang Y, Zhao L, Liu H, Yang L, Luo K. Molecular cloning and characterization of *PtrLAR3*, a gene encoding leucoanthocyanidin reductase from *Populus trichocarpa*, and its constitutive expression enhances fungal resistance in transgenic plants. J Exp Bot. 2012; 63: 2513–2524.
68. Wang Y, Chun OK, Song WO. Plasma and dietary antioxidant status as cardiovascular disease risk factors: a review of human studies. Nutrients. 2013; 5: 2969–3004.
69. Hammerbacher A, Paetz C, Wright LP, Fischer TC, Bohlmann J, Davis AJ, Fenning TM, Gershenzon J, Schmidt A. Flavan-3-ols in Norway spruce: biosynthesis, accumulation, and function in response to attack by the bark beetle-associated fungus Ceratocystis polonica. Plant Physiol. 2014; 164: 2107–2122.
70. Wang PQ, Zhang LJ, Jiang XL, Dai XL, Xu LJ, Li T, Xing DW, Li YZ, Li MZ, Gao LP, Xia T. Evolutionary and functional characterization of leucoanthocyanidin reductases from *Camellia sinensis*. Planta. 2018; 247: 139–154.
71. Koes RE, Quattrocchio F, Mol JN. The flavonoid biosynthetic pathway in plants: function and evolution. Bioessays. 1994; 16: 123–132.
72. Takos AM, Ubi BE, Robinson SP, Walker AR. Condensed tannin biosynthesis genes are regulated separately from other flavonoid biosynthesis genes in apple fruit skin. Plant Sci. 2006; 170: 487–499.
73. Tuttle, JR, Idris, AM, Brown JK, Haigler CH, Robertson D. Geminivirus-mediated gene silencing from *Cotton Leaf Crumple Virus* is enhanced by low temperature in cotton. Plant Physiol. 2008; 148: 41–

50.

Additional Files

Additional file 1: Table S1 Primers used in the experiments

Additional file 2: Figure S1 Relative expression levels of *GhCHS*, *GhANR*, *GhANR* in the developing fiber of 5 DPA, 10 DPA and 15 DPA in *GhCHSi*, *GhANRi*, *GhANRi*, *GhPDSi* transgenic cotton lines and wild cotton ZX1.

Figures

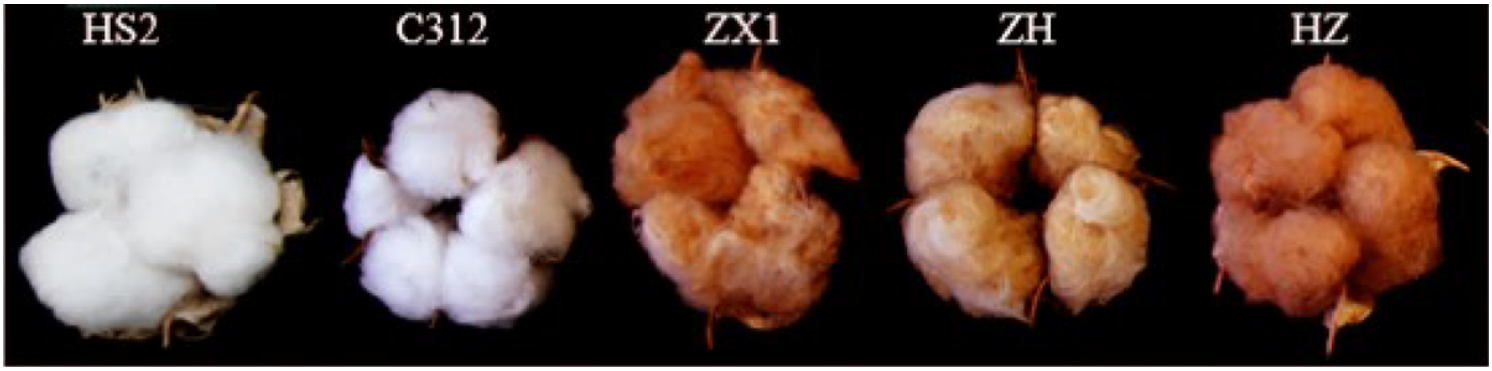

Figure 1

The phenotype of cotton bolls and fiber in HS2, C312, ZX1, ZH, HZ (G. hirsutum L.) used in the experiments.

Figure 2

Clustal W alignment of multiple amino acid sequences alignment of GhCHS, GhANR and GhLAR. Colors indicate the similarity of amino acids sequences coded by the GhCHS (A), GhANR (B), GhLAR (C).

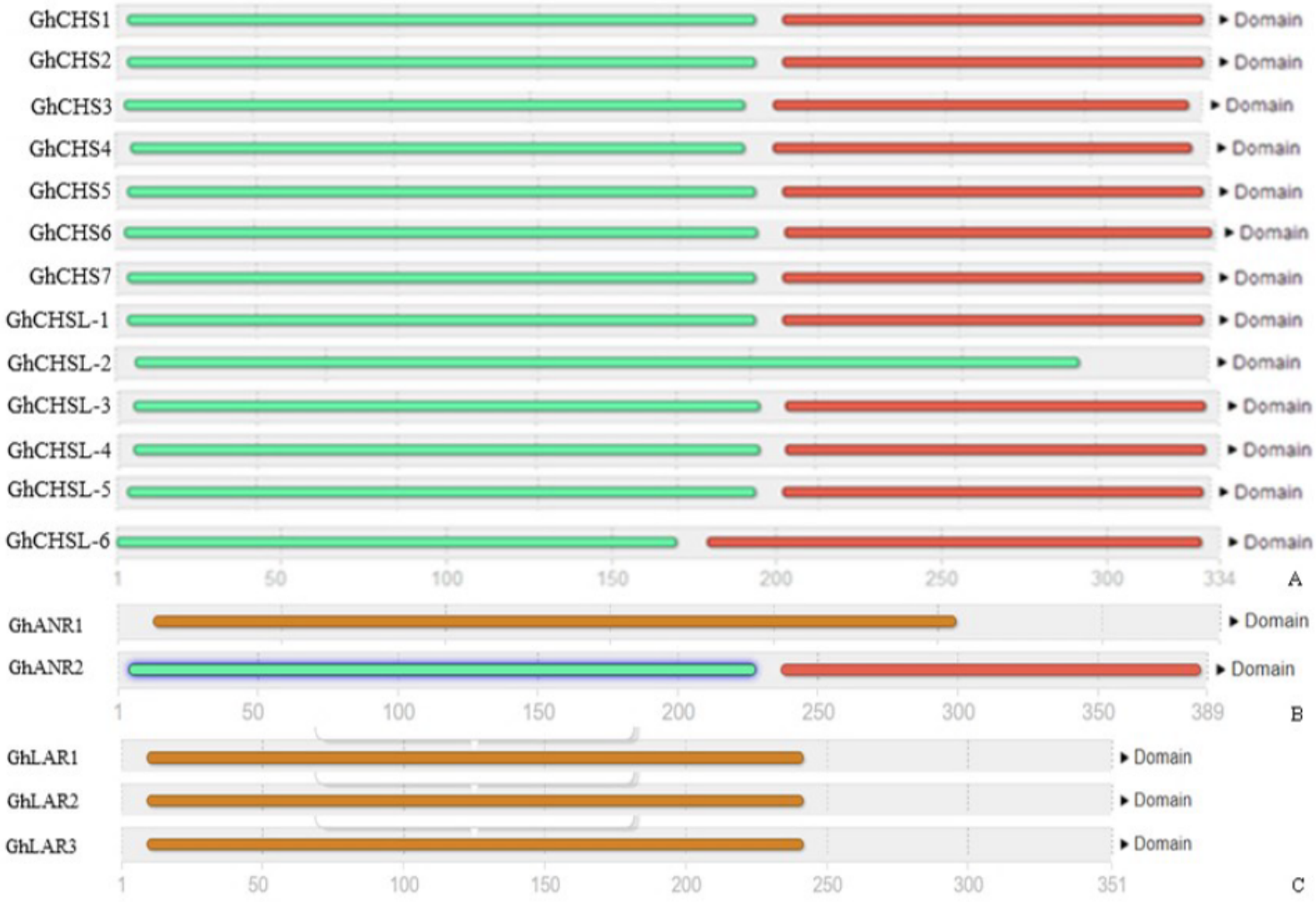

Figure 3

Structural and motif analysis of GhCHS, GhANR and GhLAR based on InterPro test

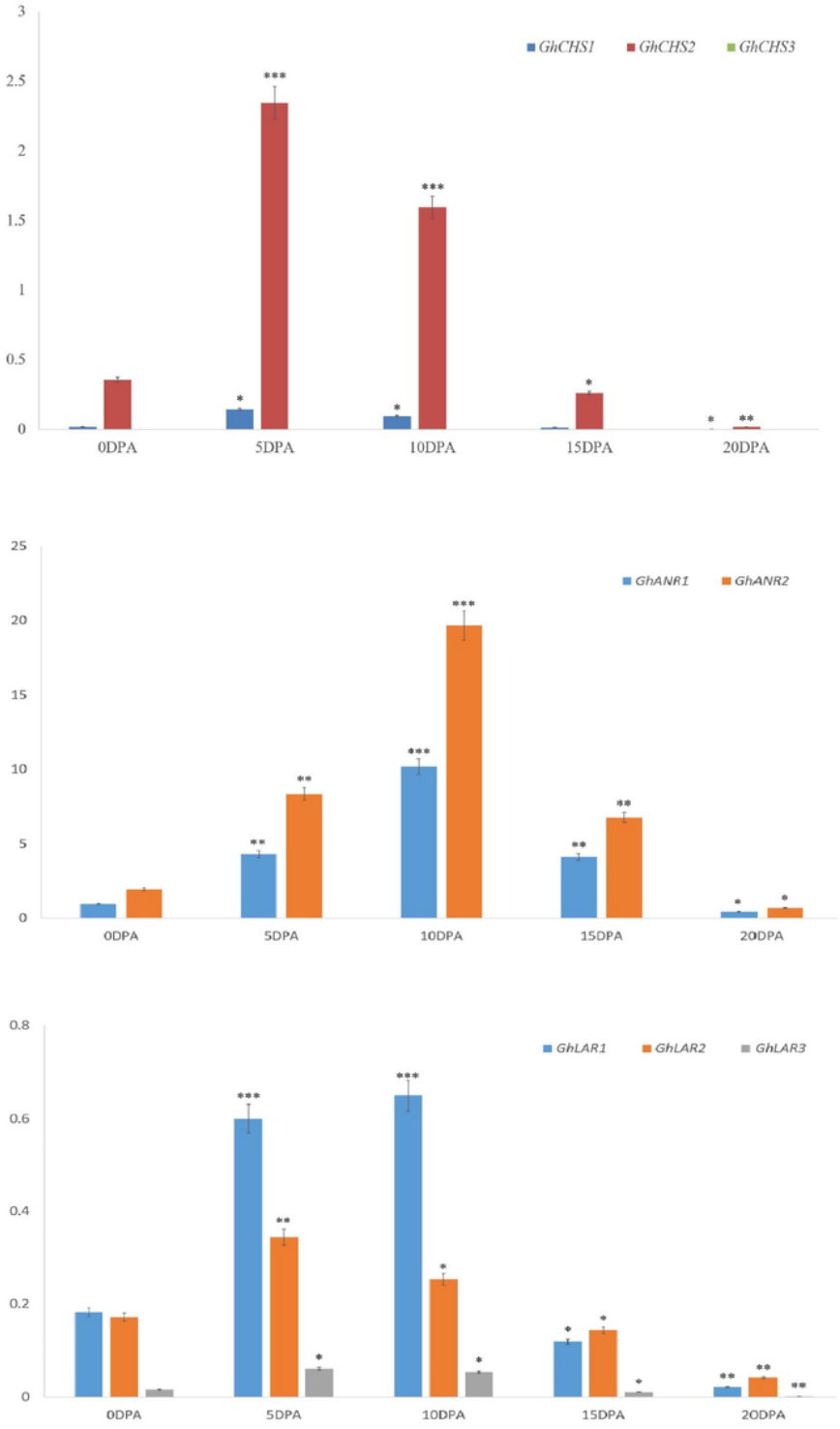

Figure 4

The expression analysis of GhCHSs, GhLARs and GhANRs in the developing fiber of 0, 5, 10, 15 and 20 DPA in natural brown cotton ZX1. Data presented in all graphs are means ± SD (n =3). (Student's t-test, **P < 0.05, **P < 0.01, ***P < 0.005, compared to 0 DPA)

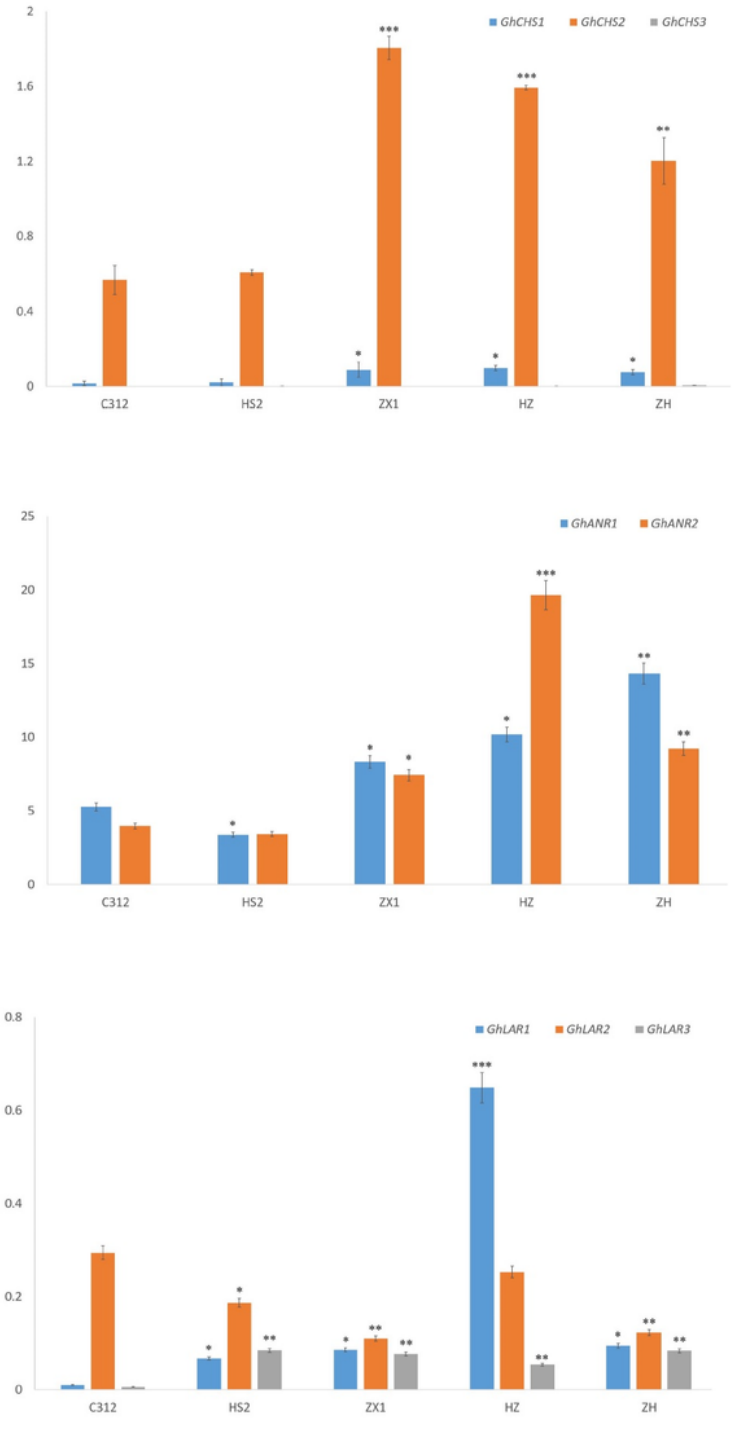

Figure 5

The expression analysis of GhCHS, GhLAR and GhANR genes in the developing fiber of 10 DPA in 5 cotton species. Data presented in all graphs are means ± SD (n =3). (Student's t-test, *P < 0.05, **P < 0.01, ***P < 0.005, compared to C312)

Figure 6

Phenotypic analysis of plant, boll and fiber in GhCHS-RNAi (GhCHSi), GhANR-RNAi (GhANRi), GhANR-RNAi (GhANRi), GhPDS-RNAi (GhPDSi), negative controls (N CK) of ZX1 and wild cotton of ZX1 plants (WT ZX1). A: the phenotype of negative controls ZX1 (N CK) and wild cotton ZX1 plants (WT ZX1). B: the transgenic plants of GhCHSi, GhANRi, GhANRi. C: the transgenic plants of GhCHSi, GhANRi, GhANRi with opening bolls. D: the phenotype of boll and fiber of GhCHSi, GhANRi, GhANRi plants.

Figure 7

The phenotypic comparison of boll and fiber in GhCHSi, GhANRi, GhANRi, GhPDSi, and different controls of white fiber C312, donor cotton of natural colored cotton (ZX1, HZ and ZH) and C312. A: the phenotypic comparison of fiber in transgenic ZX1 lines of GhCHSi (A), GhANRi (B), GhANRi (C), GhPDSi (positive control ZX1) and different controls of white fiber C312, ZX1, negative control ZX1 (N CK, transgenic ZX1 with vector-free). D: the phenotypic comparison of fiber in transgenic HZ lines of GhCHSi, GhANRi, GhANRi, and different controls (HZ with dark brown fiber). E: the phenotypic comparison of fiber in transgenic ZH lines of GhCHSi, GhANRi, GhANRi, and different controls (ZH with light brown fiber).

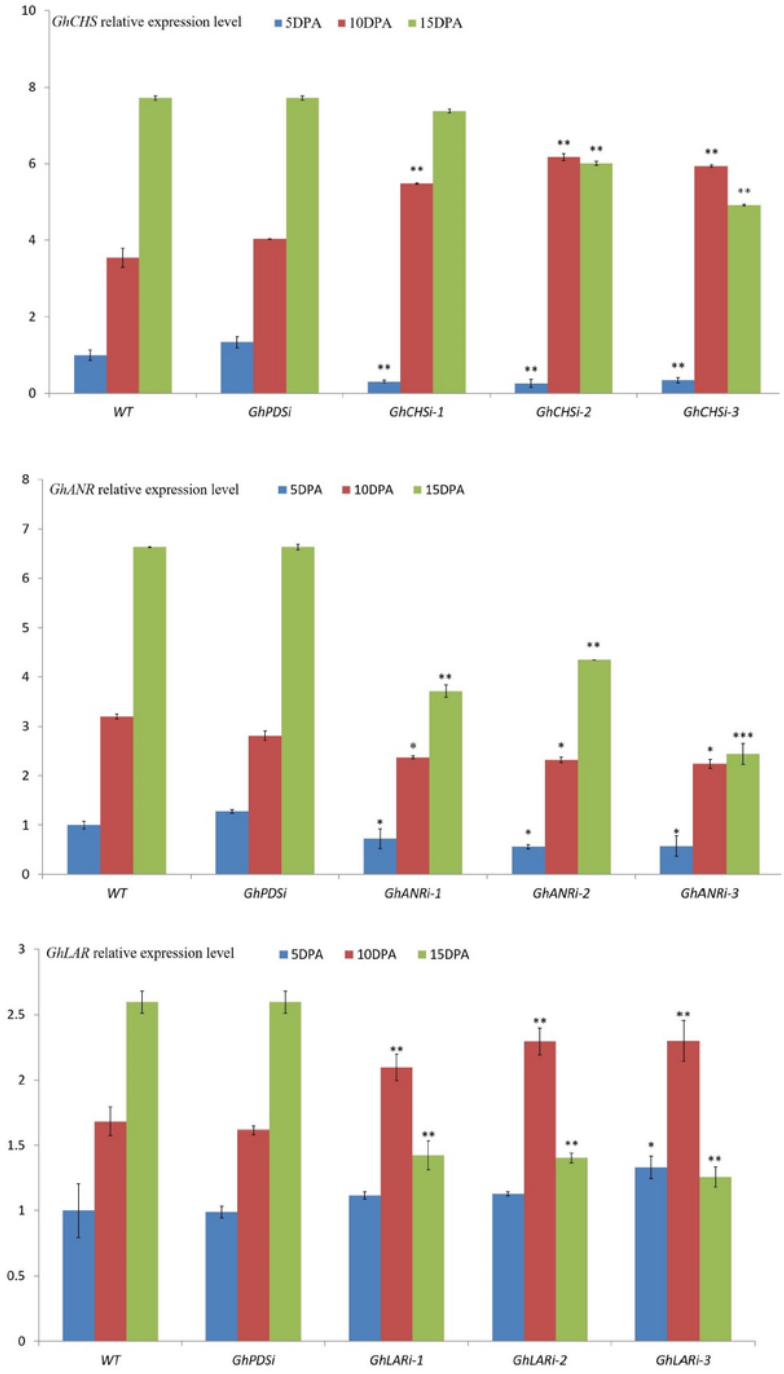

Figure 8

Relative expression levels of GhCHS, GhANR, GhANR in the developing fiber of 5 DPA, 10 DPA and 15 DPA in GhCHSi, GhANRi, GhANRi, GhPDSi transgenic cotton lines and WT ZX1. A: Relative expression analysis of GhCHS in GhCHSi, GhPDSi transgenic cotton lines and WT. B: Relative expression analysis of GhANR in GhANRi, GhPDSi transgenic cotton lines and WT. C: Relative expression analysis of GhLAR in GhLARi,

GhPDSi transgenic cotton lines and WT. Data presented in all graphs are means ± SD (n =3), (Student's t-test, *P < 0.05, **P < 0.01, ***P < 0.005, compared to WT)

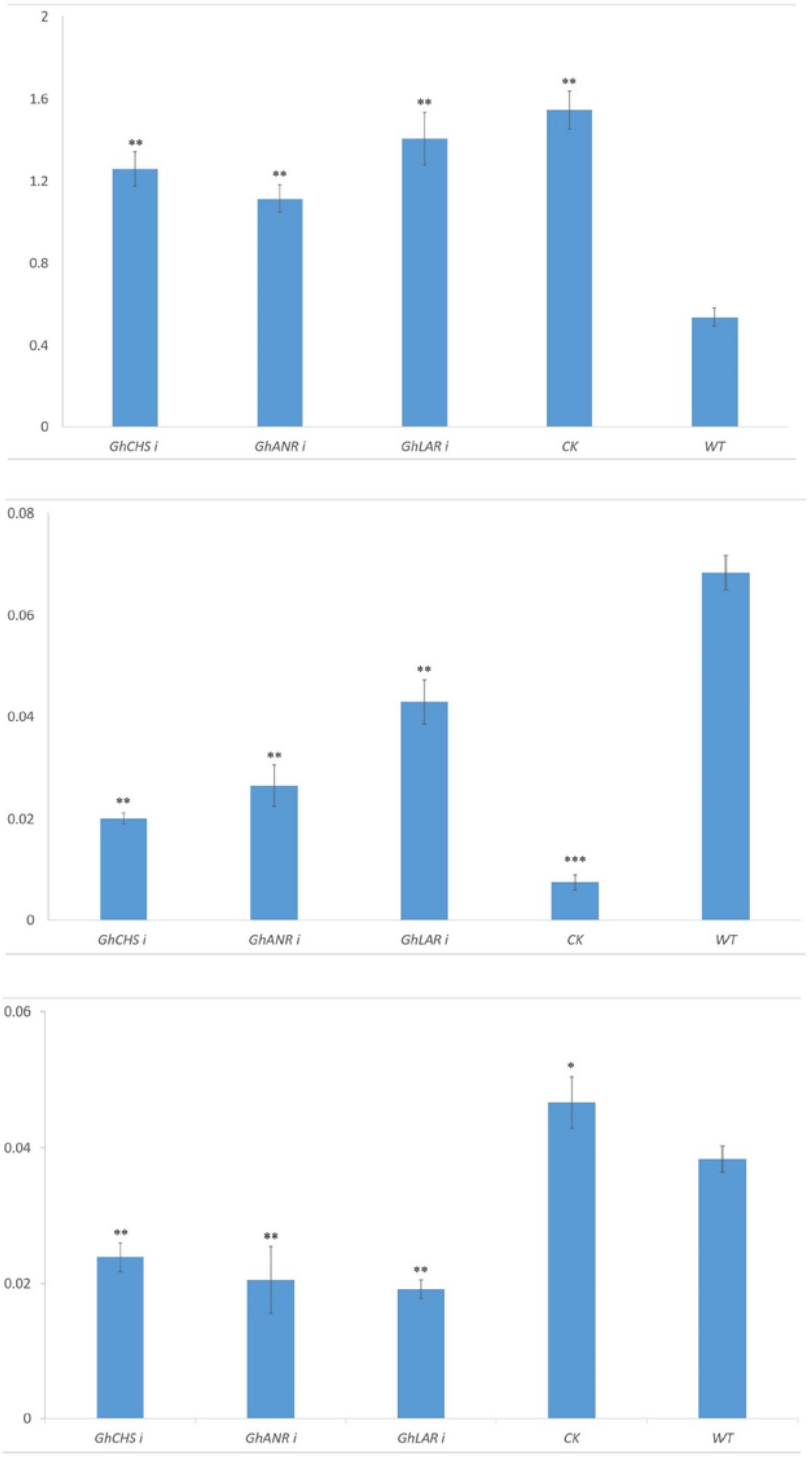

Figure 9

The anthocyanidins content in transgenic plants of GhANRi, GhLARi and GhCHSi, ZX1 (WT) and N CK ZX1. A: The anthocyanidins content in leaves of GhANRi, GhLARi and GhCHSi plants, ZX1 and N CK (Transgenic ZX1 with vector-free); B: The anthocyanidins content in cotton kernels of GhANRi, GhLARi

and GhCHSi plants, WT and N CK; C: The anthocyanidins content in fiber and seedcoat in GhANRi, GhLARi and GhCHSi plants, ZX1 and N CK.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfiles.docx