

Molecular cloning and characterization of GhERF105, a gene participated in the regulation of gland formation from cotton (*Gossypium hirsutum* L.)

Chaofeng Wu

Institute of cotton research

Hailiang Cheng

Cotton Research Institute

Shuyan Li

Cotton Research Institute

Dongyun Zuo

Cotton Research Institute

Zhongxu Lin

Huazhong Agriculture University

Youping Zhang

Cotton Research Institute

Limin Lv

Cotton Research Institute

Qiaolian Wang

Cotton Research Institute

Guoli Song (✉ songguoli@caas.cn)

Cotton Research Institute <https://orcid.org/0000-0003-3236-9286>

Research article

Keywords: Cotton (*Gossypium hirsutum* L.), GhERF105, gland formation

Posted Date: August 25th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-36598/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published on February 18th, 2021. See the published version at <https://doi.org/10.1186/s12870-021-02846-5>.

Abstract

Background

Gossypium hirsutum L. (cotton) is one of the most economically important crops globally. Cottonseed is the significant source of fiber, feed, foodstuff, oil and biofuel products. However, the utilization of cottonseed was limited by the presence of small and darkly pigmented glands that contain large amounts of gossypol, which is toxic to human beings and other non-ruminant animals. To date, there has been some progress in the pigment gland formation, but the underlying molecular mechanism of pigment gland formation was still complicated and unclear.

Results

In this study, we identified an AP2/ERF transcription factor named GhERF105 (Gh_A12G1784), which was involved in the regulation of gland pigmentation, from comparative transcriptome analysis of the leaf of two pairs of glanded and glandless accessions, which are CCRI12 and CCRI12XW, L7 and L7XW. It encoded an ERF protein localized in the nucleus with transcriptional activation activity containing a conserved AP2 domain, and showed the high expression in glanded cotton accessions that contained much gossypol. Virus-induced gene silencing against *GhERF105* caused the dramatic reduction in the number of glands and significantly lowered levels of gossypol in cotton leaves. *GhERF105* showed the patterns of spatiotemporal and inducible expression in the glanded plants.

Conclusions

These results suggest that *GhERF105* participates in the pigment gland formation and gossypol biosynthesis in partial tissue of glanded plant. It also provides a potential molecular basis to generate 'glandless-seed' and 'glanded-plant' cotton cultivar.

Background

Cotton (*Gossypium* spp.) is a globally appreciated crop for its economic value of the textile fiber, feed, foodstuff, oil and biofuel products in the world [1–2]. There are approximately 50 species in the *Gossypium* genus, four of which are cultivated in agriculture, including two diploid cottons (*G. herbaceum* and *G. arboreum*) and two allotetraploid cottons (*G. hirsutum* and *G. barbadense*) [3–5]. *G. hirsutum*, the upland cotton, is most widely grown today and dominates world cotton commerce with more than 90% of the annual cotton fiber production [5–6]. However, the potential of nutrition sources of cottonseed cannot be used sufficiently due to the presence of gossypol stored mainly in the small darkly pigmented lysigenous glands in plants and seeds, which is toxic to human beings and other non-ruminant animals [7–8]. Gossypol, as phytoalexin, is a yellowish phenolic compound that serves as a protective function against various pests, diseases and abiotic stresses in certain species of cotton plants of the family

Malvaceae [9–11]. Therefore, developing cotton with low-gossypol seeds and high-gossypol plants has become an interesting area of cotton breeding for researchers.

The pigment glands, which are also called ‘gossypol glands’, ‘internal glands’ and ‘black glands’ located in certain of the subepidermal layer of hollow organs in many parts of the plant, originate from a cluster of cells in the ground meristem, which differ from other cells in that they have a high-density gossypol and related terpenoids [7]. Research on the molecular genetic mechanisms of pigment gland in the cotton plant began in lines of ‘Hopi Moencopi’ in the 1950s [12–14]. So far, the research of cotton inheritance has indicated that the gland formation is controlled by a combination of at least six independent loci such as gl_1 , gl_2 , gl_3 , gl_4 , gl_5 and gl_6 , the different combinations of dominant(G) and recessive (g) alleles modulate gland formation in different organs [14–19]. The completely glandless phenotype was controlled by two pairs of duplicate homozygous recessive genes ($gl_2gl_2gl_3gl_3$) in the allotetraploid *G.hirsutum* [13–14], while the dominant alleles (Gl_2 , Gl_2 , Gl_3 , Gl_3) in any combination produced the glanded phenotype with variable distribution in different organs [14, 20]. In 1965, the gl_2 and gl_3 gene were located on chromosome (chr.) A_112 and D_112 of *G.hirsutum*, respectively [16, 21–22]. Alleles gl_4 and gl_5 decrease the number of glands while gl_6 have the weaker effects on gland formation compared with gl_1 [23–24]. Subsequently, gl_2^{arb} , gl_2^b , gl_3^{dav} , gl_3^{thur} , gl_3^{rai} , gl_3^b [7], Gl_2^s [25], Gl_2^e [26–27] and Gl_2^b [28] related to pigment gland formation were identified. Among them, Gl_2^e is the most critical gene that controls glandless character of the whole plant. A single completely dominant glandless *G. barbadense* mutant(Gl_2^e) named ‘Bahtim 110’ (*G. barbadense* L), which is a dominant allele of Gl_2 that shows epistatic effect on Gl_3 , was originally discovered in Egypt by the irradiation mutagenesis of the sea-island cotton ‘Giza 45’ seeds with ^{32}P , and could efficiently inhibit the formation of pigment gland[29–32]. Since then, several genes for gland formation have been discovered gradually by researcher. In 2016, *GoPGF* gene(Gossypium Pigment Gland Formation gene),which encodes a basic helix-loop-helix transcription factor as the critical gene, was identified through map-based cloning approach and located on chromosome A_112 [33–34]. *CGF3* (Cotton Gland Formation), which is identical to *GoPGF* gene, not only controls the gland morphogenesis directly, but also regulates gossypol biosynthesis indirectly [35]. *CGP1* (Cotton Gland Pigmentation 1), which interacted with *GoPGF*, was identified through comparative transcriptome analysis of glanded and glandless cotton accessions and was involved in the regulation of gossypol biosynthesis but not gland formation [36]. In addition, the novel RanBP2 zinc finger protein (ZFP) and GauGRAS1, which played the roles in the development of the cotton gland, were identified using suppression subtractive hybridization (SSH) from upland cotton ‘Xiangmian 18’ [9, 37–39]. During the past three decades, there has been some progress in the molecular mechanism of gland formation and the relationship between gossypol and pigment gland. However, the specific mechanism of pigment gland formation remains complicated and still unclear up to now.

Here, we identified an Ethylene Response Factor named GhERF105, which was involved in the regulation of gland pigmentation, from a comparative transcriptome analysis of the leaf of two pairs of glanded and glandless allotetraploid cotton accessions, which are CCRI12 and CCRI12XW, L7 and L7XW(Fig.S1).

It encoded an ERF protein localized in the nucleus with transcriptional activation activity containing a conserved AP2 domain and showed the high expression in glanded cotton accessions that contained much gossypol. Silencing of GhERF105 by VIGS not only resulted in the drastic reduction of gland, but also decreased the accumulation of gossypol in the leaves of the treated plants. Moreover, GhERF105 showed a temporal and spatial pattern of expression in various hollow organs of glanded and glandless cotton plants, including Cotyledon, Hypocotyl, Petiole, True leaf and Stem, and demonstrated the inducible expression under ethylene treatment. In addition, *GhERF105*, *CGF*, *CGP1* and *GoPGF* genes were highly expressed in the leaves and stems in glanded CCRI12 and L7 but were indeed substantially lower expression in CCRI12XW, CCRI12YW and L7XW.

These results provide a reference for the comprehensive analysis of the molecular mechanism of formation of gland and gossypol biosynthesis in cotton. However, the diversity of gland trait inheritance indicates the regulation complexity of gland formation. Further studies are needed to better understand the molecular mechanisms underlying gland development.

Results

Sequence analysis of the full-length *GhERF105* gene

In this study, we identified an Ethylene Response Factor named GhERF105 (Gh_A12G1784) through a comparative transcriptome analysis of the leaves of two pairs of glanded and glandless accessions, which are CCRI12 and CCRI12XW, L7 and L7XW. The *GhERF105* gene, which was cloned from the leaves of CCRI12 is 711bp in length with an open reading frame (ORF) of containing initial code (ATG) and terminal code (TAA) with no intron (Fig.1). The predicted protein comprised of 236 amino acids with relative molecular weight of 26.3kDa and isoelectric point of 7.72 containing an ERF conserved DNA binding domain. The cotton GhERF105 belongs to the largest AP2/ERF family of transcription factors that plays an important role in plant development and environmental stress responses, as well as hormone signaling and pathogen defense [40-42].

The expression analysis of *GhERF105* gene in many cotton accessions

Because *GhERF105* gene was identified from the comparative transcriptome analysis of the leaf of two pairs of Near Iso-genic Lines (NILs) with glanded and glandless phenotype. Therefore, the expression levels of *GhERF105* were analyzed in two pairs of Near Iso-genic Lines (NILs) and other cotton accessions, the results show that *GhERF105* was highly expressed in the leaves and stem of glanded *G.hirsutum*. (CCRI12, L7 and TM-1) but was indeed substantially lower expression in CCRI12XW, L7XW and CCRI12YW (Fig.2). Based on the different expression pattern of *GhERF105* in partial tissues of six cotton accessions, *GhERF105* may be related to the formation of glands. However, its function and regulatory mechanism in pigment gland development need further be investigated using VIGS technology and other technology.

Silencing of *GhERF105* reduced gland formation and gossypol biosynthesis

Here, In order to further ascertain the function of the *GhERF105* during pigment gland formation. *Agrobacterium*-mediated VIGS systems was constructed using a TRV-based VIGS vector for silencing *GhPDS* and *GhERF105* genes in the cotton seedlings. The results were as follows: (1) Silencing the expression of the endogenous phytoene desaturases gene (PDS), which is commonly used marker gene for VIGS, causes loss of chlorophyll and carotenoids [43]. A photobleaching phenotype in cotton plants infiltrated with *GhPDS*-expressing agrobacteria was observed 14-21 days after infiltration in true leaves, compared to the leaves in plants infiltrated with negative gene-expressing agrobacteria (Fig.3a). (2) To assess its function, we cloned the 289bp fragment of *GhERF105* from CCRI12 plant and inserted it into pTRV2 for virus-induced gene silencing (VIGS) to suppress the expression of endogenous in cultivated glanded allotetraploid cottons. Compared with that in the untreated CCRI12 as the negative control (Fig.3b1-b2), The *GhERF105*-silenced CCRI12 plants exhibited the dramatic reduction in gland numbers in the new true leaf of 14-21d after infiltration (Fig.3 b3-b6). The transcript levels of *GhERF105* in pTRV-*GhERF105* leaves were prominently lower than those in the untreated CCRI12 but still higher than those in the untreated CCRI12XW (Fig.3c). However, the veins of the new emerging true leaves had fewer dotted glands and the stems had thickly dotted glands (Fig.3 b5-b6, Fig.S2). These data suggested that *GhERF105* regulated the formation of glands in leaf but not stem, in contrast to *GoPGF*, which results in glandless phenotype in all tissues, including the leaves and stems [29]. (3) we conducted HPLC analysis to measure the level of gossypol in the leaves. Gossypol content was reduced by over 78% in the *GhERF105*-silenced leaves compared with the untreated CCRI12 leaves but still higher than those in the untreated CCRI12XW (Fig.3d). In all, the results suggest that *GhERF105* may be involved in the pigment gland formation and gossypol biosynthesis.

Spatiotemporal expression analysis of *GhERF105* gene

The pigment glands are located on the surfaces of the stem, leaves, sepals, petals, and stigma [17]. *GhERF105* gene was associated with the development of cotton pigment gland. Therefore, the transcription level of *GhERF105* gene was detected by qRT-PCR in different organs of gland development in glanded and glandless cotton accessions. The result showed the mRNA levels in Cotyledon, Hypocotyl, Petiole, True leaf and Stem of the gland plant were increased to 3.5, 10.5, 15, 8.7 and 4.0 folds of that in glandless plant, respectively. That is to say, the mRNA levels of *GhERF105* in the above organs of the glanded plants was higher than that in the glandless plants. At the same time, the expression level of *GhERF105* was highest in the leaf of glanded plants but wasn't significant differences between the leaf and other organs of glandless plants (Fig.4). Therefore, the *GhERF105* gene is of a highly different expression pattern between the glanded and glandless cotton plants in pigment gland formation.

Nuclear Localization and Revealed Transcription Activity of *GhERF105* protein

The green fluorescent protein (GFP) reporter, which is a vital marker for protein subcellular localization, showed a very strong fluorescence signal under the control of the constitutive CaMV35S promoter, and the signal was uniformly and diffusely distributed throughout the cell. Based on functional annotation information, *GhERF105* is believed to act as a transcriptional factor. Therefore, the nuclear localization

should be essential for the function of GhERF105. To test this possibility, the coding sequence (CDS) of GhERF105 was fused to the green fluorescent protein (GFP) reporter gene. After introducing the construct (Fig. 5a,S3) into the tobacco cells by agro infiltration, GhERF105-GFP, the transcription factor fused to GFP, was expressed transiently and located exclusively in the nucleus of tobacco epidermal cells (Fig.5b). This confirms that GhERF105-GFP is a nuclear localized protein that is mainly involved in nuclear transport.

The yeast strains transformed with the pGBKT7-GhERF105 were able to grow blue colonies on the selected medium SD/-Trp/-X-a-gal while those strains with empty vector pGBKT7 could grow white colonies (Fig.6). This result indicated that pGBKT7-GhERF105 is of the transcriptional activity, implicating a role of GhERF105 as a transcription activator.

Expression pattern of *GhERF105* gene in cotton under ethylene treatment

The ERFs, which are important plant-specific transcription factors in the ethylene signal transduction pathway, have been shown to play an important regulatory role in modulating the expression of specific stress-related genes [44-46]. Ethylene interacting with other plant hormone, regulated the programmed expression of pathogenesis-related (PR) genes in the ethylene-mediated signaling pathways[47].

Programmed cell death (PCD) plays an important role during the development of pigment glands in *Gossypium hirsutum* leaf tissue [48]. Ethylene, which is the upstream signal molecules during PCD process, mediates the PCD signal by ROS [49]. Therefore, It is meaningful to investigate the expression pattern of *GhERF105* gene in response to stress hormone ethylene stimuli. In this study, qRT-PCR analysis was employed to detect *GhERF105* expression level changes in leaves at different time after ethylene treatment. Compared to that in the water-treated plants, The *GhERF105* mRNA was rapidly accumulated and peaked within 8h, and then declined to the original level in the ethylene-treated plants, The result suggested that the level of the mRNAs of *GhERF105* gene was inducible at the early stage of ethylene treatment and maintained the high level from 6h to 10h by the stress hormone ethylene in leaf (Fig.7). However, the expression pattern of *GhERF105* did not positively correlate with the length of time after ethylene treatment. The result indicates that the expression of *GhERF105* is was characteristically responsive to ethylene treatment at the transcriptional level and *GhERF105* may be related to ethylene signal transduction pathways or defense/stress signaling pathways. At the same time, It is tempting to speculate that gland formation and gossypol synthesis in cotton may be induced and regulated directly or indirectly by ethylene.

Expression Patterns of genes involved in gland formation

To see whether there is the relationship between *GhERF105* and other genes (such as *CGF1*, *CGF2*, *CGP1* and *GoPGF*) that are involved in the gland formation [16, 21, 29]. The expression levels of *GhERF105*, *CGF1*, *CGF2*, *CGP1* and *GoPGF* were analyzed in the leaf and stem of five cotton accessions by qRT-PCR, including glanded *G. hirsutum*(CCRI12, TM-1 and T582), dominant glandless *CCR12XW* and recessive glandless CCRI12YW. Results obtained from qRT-PCR analysis confirmed that *GhERF105*, *CGF1*, *CGF2*, *CGP1* and *GoPGF* were highly expressed in the leaves and stems in glanded CCRI12 and TM-1 but were

indeed substantially lower expression in CCRI12XW and CCRI12YW (Fig.8). In addition, we also observed that the expression level of the genes was significantly higher in the leaves than in the stems for *G. hirsutum* (T582)(Fig.9). The results showed that *GhERF105* has the similar expression pattern as *GoPGF*, *CGF1*, *CGF2* and *CGP1* in some cotton accessions. This results further concluded that *GhERF105* was associated with cotton pigment gland development and indicated that the relationship was unknown between *GhERF105* gene and *GoPGF*, *CGF1*, *CGF2* and *CGP1* gene in gland formation.

Discussion

To date, developing cotton varieties, which produce low-gossypol seeds and high-gossypol plants, has become an important topic of cotton breeding for researchers. Therefore, it is very significant to understand the molecular mechanism of the pigment gland formation and the relationship between gossypol and gland in cotton.

In the recent years, the considerable efforts have been made by researchers to accumulate knowledge and to identify a series of genes related to pigment gland formation and gossypol synthesis. *GoPGF/CGF3*, which plays the most critical and direct role in gland development, independently regulates gland morphogenesis and indirectly affect gossypol biosynthesis possibly through regulating the expression of gossypol-related genes through binding to the G-box motif by yeast one-hybrid assays [34]. *CGF1* shows similar functions to *CGF3*, and *CGF2* regulates the density of pigment glands [35]. Silencing *GoPGF* results in the apparent absence of glands in all tissues of gland cotton and leads to an almost complete lack of gossypol [34–35]. *CGP1* modulates gossypol accumulation but not gland morphogenesis, Knockout of *CGP1* by CRISPR/Cas9 and VIGS produces a strong reduction in gossypol levels [36]. Silencing *GauGRAS1* by VIGS leads to glandless stems and petiole and didn't change the form of glands in the leaves in *G. australe*. Moreover, the gossypol content in the stem of the *GauGRAS1*-silenced plants was significantly reduced [38]. However, the molecular mechanism for pigment gland formation remains complicated and unclear, which has limited progress in low-gossypol breeding of cotton. Therefore, the further exploring of molecular mechanism for gland formation may facilitate the genetic improvement of cotton.

Here, this study provides several evidences that *GhERF105* gene was associated with gland formation of partial tissues in the glanded plant. First, *GhERF105* gene was identified through the comparative transcriptome analysis of the leaf of two pairs of glanded and glandless cotton accessions. Second, *GhERF105* was highly expressed in the glanded accession. Third, *GhERF105* knockout via VIGS markedly resulted in the drastic reduction of visible pigmented glands and decreases the content of the gossypol in the leaves but didn't change the density of gland on the stem of the cotton. In addition, the expression pattern of *GhERF105* was similar with that of another known genes related to gland development (such as *GoPGF* and *CGF*) in some glanded and glandless accessions [Fig. 7]. These findings further indicated that *GhERF105* may be involved the gland formation in cotton. Nevertheless, the regulatory mechanism on the pigment gland was somewhat different between *GhERF105* and *GauGRAS1*.

Ma et al. (2016) have confirmed that *GoPGF* can bind to G-box motif by Yeast one-hybrid assays [34]. The promoter region of *GhERF105* includes G-box cis-acting elements. It is therefore tempting to speculate that *GoPGF* likely regulates the expression of *GhERF105* by binding to G-box cis-acting elements in the nucleus and modulate the expression of gossypol-related genes by binding to the related cis-acting elements of their promoter directly and indirectly (Fig. 10), the speculation will be needed further be verified by the results of related experiments.

In conclusion, the cloning and characterization of *GhERF105* both provide new information to study the molecular mechanism of gland formation and its functions in upland cotton.

Conclusions

Based on a comparative transcriptome analysis of the leaf of two pairs of glanded and glandless cotton accessions, we identified an Ethylene Response Factor named GhERF105 that is involved in the regulation of gland pigmentation, The *GhERF105* cDNA, which was cloned from the leaves of CCRI12 seedlings, is 711 base pair (bp) in length with an open reading frame (ORF) of containing initial code (ATG) and terminal code (TAA) with no intron, The predicted protein comprised of 236 amino acids with relative molecular weight of 57.6 kDa and isoelectric point of 5.15 containing a ERF conserved DNA binding domain. The cotton GhERF105 belongs to the largest AP2/ERF family of regulatory transcription factors. The gene was differentially expressed at different organs in glanded and glandless cotton accessions. Silencing of *GhERF105* by VIGS not only resulted in the drastic reduction of the number of gland, but also decreased the accumulation of gossypol in the leaves of treated plants. GhERF105 is located in the nucleus with transcriptional activation activity and induced by ethylene. The results suggested that the novel *GhERF105* may participate in the regulation of the pigment gland and gossypol biosynthesis, as well as hormone signaling and pathogen defense.

Taken together, the cloning and functional assessment of *GhERF105* will open novel opportunities to discover the mechanism of gland formation in cotton. These results in this paper will facilitate further the improvement of cotton varieties with both glandless seeds and glanded plants through genetic engineering.

Methods

Plant Material and growth condition

Accession of CCRI12, CCRI12XW, CCRI12YW, L7, L7XW, TM-1, T582, were gifted by Cotton Research Institute, the Chinese Academy of Agricultural Sciences (CAAS) (Anyang, China), CCRI12 (China Cotton Research Institute 12) and L7 (LiaoMian 7) are Upland cotton cultivars with dark-colored pigment glands and high content of gossypol in both plants and seeds. while CCRI12XW, CCRI12YW, and L7XW, which have glandless and low gossypol content in both seeds and plants, are dominant glandless near isogenic lines (NILs) that differ nearly only in the gland trait of CCRI12 and L7, respectively [50]. 'TM-1', which is

widely used as a genetic standard, is the glanded accession of the seeds and the whole plant, 'T582' is an accession with glandless-stem and glanded-leaf of plant traits. All materials were stored by self-crossing for several years in our lab.

The seeds were immersed in water and followed by germination in a high humidity environment at 28°C in the dark for 2d. Well-germinated seeds were subsequently planted in 0.3 litres pots of 7cm diameter with one seed per pot in a commercially available sand/soil/fertilizer mix and grown for two to three weeks at 28°C (16h light and 8h dark) with LED lamps (Oppl lighting Zhongshan China) in a greenhouse.

Extraction of total RNAs

Samples from different tissues of the cotton plants, including Cotyledon, Hypocotyl, Petiole, True leaf and Stem of one or many different gland accessions served as the source of total RNA were immediately frozen in liquid nitrogen and stored at -80°C. For each sample, total RNAs were isolated from 100 mg of leaf ground with liquid nitrogen using the RNAPrep Plant RNA kit (polysaccharides&polyphenolics-rich) (TIANGEN BIOTECH (BEIJING)CO., LTD) according to the manufacturer's instructions. The quantity and purity of RNAs were assessed according an absorbance ratio of OD_{260/280}(1.9-2.1) using a NanoDrop One^C Microvolume UV-Vis Spectrophotometer with Wi-Fi (Thermo Fisher Scientific Inc., Waltham, MA, USA) ultraviolet spectrophotometer, and was confirmed using 1.0% (w/v) denatured formaldehyde agarose gel electrophoresis to investigate its integrality.

Synthesis of the first-strand cDNA

RNA was reversely transcribed into 1st strand cDNA in a 20µL reaction volume using the PrimeScriptTM1stStrand cDNA Synthesis Kit (TaKaRa Bio, Dalian, China) following the manufacturer's protocol of Reverse Transcription System. Firstly, two micrograms(2µg) of total RNA was mixed with 1.0µL Oligo dT Primer(50µM), 1.0µL dNTP mixture(10mM each), then RNase free ddH₂O was added to make the whole reaction volume up to 10µL, then, A total 10µL reaction volume was incubated at 65°C for 5min and placed on ice for 2 min to denature probable RNA secondary structure. Secondly, the first-strand cDNA synthesis mixture was prepared by adding following components to the above 10µL reaction volume in the indicated order, 4µL 5xPrimeScript II Buffer, 0.5µL RNase Inhibitor(40U/µL), 4.5µL RNase free ddH₂O, and 1µL Primescript II RTase(200U/µL). The first-strand cDNA synthesis mixture was incubated at 30°C for 10 min, 42°C for 60min and terminated at 95°C for 5 min.

Gene Expression Analysis by Quantitative Real-Time PCR

The RNA sequencing samples that were isolated were used to perform real-time quantitative (qRT-PCR) analysis using the ABI Quantstudio 5 Detection System (Applied Biosystems, Carlsbad, CA). Actin (GenBank accession numbers: AY305733) was used as reference gene. The gene-specific primers with about 215bp product size were designed using the Primer 5.0 software or online in NCBI website (https://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi?LINK_LOC=BlastHome) and listed in Table

11. The specificity of each primer set was validated by melt-curve analysis, and the efficiency was calculated by analyzing the standard curves with a tenfold cDNA dilution series (Bustin et al., 2009). The 20 μ L qPCR experiment was carried out on the ABI Quantstudio 5 Detection System with TB Green Premix Ex TaqTM (Tli RNaseH Plus) (TaKaRa Bio, Dalian, China). The reaction volume contains 0.5 μ L of each primer (10 μ M), 0.4 μ L ROX Reference Dye II (50x), 1 μ L above synthesized cDNA template, and 7.6 μ L of sterilized ddH₂O. The qPCR thermal cycling conditions were 95°C for 5 min to pre-denature cDNA template; 40 amplification cycles of 95°C for 5s, 55°C for 30s, and 72°C for 30s; and followed by 15s at 95°C, 1 min at 60°C, and 15 s at 95°C. Each sample was run in triplicate, each biological replicate was assessed three times. The relative expression level of the genes were calculated according to the $2^{-\Delta\Delta CT}$ method [51]. For the reference gene used in this experiment, their geometrical mean was operated at first, and then the relative transcript level of target gene was calculated following the method of one reference gene. Results were generally expressed mean \pm standard error (ER) from values three independent tests.

PCR Procedure

All primers were synthesized by Sangon Biotech (Shanghai) Co., Ltd. and GENEWIZ (Suzhou) Co., Ltd. high-fidelity DNA polymerase, dNTPs and other reagents were supplied by TaKaRa Bio, (Dalian) Co., Ltd. PCR amplification of the *GhERF105* gene was performed in a reaction volume of 15 μ L containing 1 μ L of template cDNA, 1.2 μ L of dNTP Mixture (2.5mM), 3 μ L of 5 \times PrimeSTAR GXL Buffer, 0.3 μ L of Prime STAR GXL DNA Polymerase, 0.4 μ L of forward primer and 0.4 μ L of reverse primer (10 μ M for each), and 8.7 μ L ddH₂O. The amplification product was achieved using the following profile: 5 min at 98°C; 35 cycles of 10 s at 98 °C, 15s at 55°C, 1 min at 68°C and a final cycle of 5 min at 68°C; hold at 10°C. The PCR products were then purified following the instructions in the QIAquick PCR Purification Kit (250) (Qiagen, Düsseldorf, Germany) and eluted in a final sample volume of 35 μ L of Qiagen EB buffer. three microliters of each PCR product were assessed by size on a 1% agarose gel to select fragments in the range of 700bp \pm 50 bp.

VIGS procedure

The VIGS (Virus-induced gene silencing) vector tobacco rattle virus (TRV) invades a wide range of hosts and is able to spread vigorously throughout the entire plant but produces only mild symptoms [52]. Therefore, VIGS system has been proven to be a powerful tool in elucidating gene function and functional genomics in cotton [10, 34, 35-38, 53, 54]. To knockdown the expression of *GhERF105*, The pTRV-VIGS vectors were constructed using a previously published method [53, 55]. Briefly, cDNA fragments of cotton PDS (GhPDS1, 327bp, GenBank accession numbers: HQ441184) and Pigment gland formation *GhERF105* (337bp) were amplified using Prime STAR GXL DNA Polymerase (TaKaRa) from CCRI12 by PCR with gene specific primers (listed in the table 2). The resulting products were cloned into pTRV2 with *Bam*HI and *Kpn*I to produce recombinant vectors named pTRV2::PDS and pTRV2::GhERF105, respectively. These recombinant vectors and the empty vector (pTRV2::00) were then introduced into the Agrobacterium strain GV3101 (Weidi Bio, Shanghai, China) by Heat shock method, For the VIGS assay, the

transformed *Agrobacterium* colonies containing pTRV1 and pTRV2-GhPDS, pTRV2-GhERF105 were grown overnight at 28°C in an antibiotic selection medium containing rifampicin, Gentamicin and kanamycin 50mg/ml, and suspended in the solutions (10mM 2-(N-morpholino) ethane sulfonic acid, 10mM MgCl₂ and 400µM acetosyringone (AS)) to the final optical densities as OD values of 1.5 at 600nm and then left at 25°C for 4h without shaking in the dark. Before infiltration, *Agrobacterium* cultures containing pTRV1 and pTRV2 or its derivatives were mixed in 1:1 ratios. Seedlings with the fully expanded cotyledons but without a visible true leaf of CCRI12 were infiltrated by inserting the *Agrobacterium* suspension containing pTRV1 and pTRV2, pTRV2-GhPDS, pTRV2-GhERF105 into the cotyledons via a syringe. Plants were grown in the pots at 25°C in a growth chamber under a 16h light/8h dark photoperiod with 70% humidity. To analyze silencing efficiency, RNA was extracted and qRT-PCR was performed. The *Actin* (GenBank accession numbers: AY305733) and *GhERF105* was amplified as reference gene and target gene, respectively [56]. Leaves were numbered sequentially such that number 1 refers to the first true leaf initiated after the cotyledons. In this study, leaves 2-3 were investigated and collectively referred to as total foliage [57].

Gossypol detection and analysis

The total gossypol concentration in the leaves from CCRI12, GhERF105-silenced CCRI12 and CCRI12XW plants was determined by high-performance liquid chromatography (HPLC) (Agilent 1100, Agilent, Santa Clara USA). Each 100 mg plant sample, which was freeze-dried and ground into powder with liquid nitrogen. was dissolved with 2ml leaf extraction (acetonitrile/water/phosphoric acid=80:20:0.1). The leaf extraction was centrifuged at 10000rpm for 10 min and then the supernatant was carefully transferred into a new EP tube at room temperature. The eluent was filtered through a 0.45µm nylon filter into a vial for HPLC analysis with Agilent Zorbax Eclipse Plus C18 analytical column (250mm×4.6mm, 5micron). The sample was analyzed at a wavelength of 235nm. The concentration was calculated using Agilent 1100 system by comparing to the gossypol standard curve. A gossypol reference standard was purchased from Sigma Chemical Co. Ltd.

Subcellular Localization of GhERF105 Protein

To study the subcellular localization of GhERF105 protein, the coding regions of GhERF105 was amplified with stop codon removed by Primers listed in Table 3, which contained a *Xba*I and *Sma*I site (underlined) through polymerase chain reaction (PCR), The resulting fragments were cloned between the *Xba*I and *Sma*I site of the transient expression pBI121-GFP vector, which harbors an ORF encoding the green fluorescent protein (GFP) under the control of the CaMV35S promoter, and construct the recombinant plasmid *p35S-GhERF105-GFP*. p35S-GFP was used as positive control. The plasmids of GFP-GhERF105 and GFP were then introduced into tobacco leaves (*Nicotiana benthamiana*) respectively via *Agrobacterium*-mediated transformation and incubated at 25°C under light for 48-72h. The green fluorescence signals were observed and the localization of the fusion protein was determined using a confocal laser scanning microscope (Leica TCS SP8, Germany).

Transactivation Activity Assay of GhERF105 Protein

To study the transactivation activity of GhERF105 protein, GhERF105 cDNA was amplified with by Primers listed in Table 4 and cloned into the *EcoRI* and *NotI* sites of pGBKT7 vector to generate pGBKT7-GhERF105 construct. This plasmid with empty vector control was then transformed into yeast strain AH109 to analysis the transactivation activity. Yeast transformants with OD600 of 0.1, 0.01 and 0.001 were plated on the selective media, SD/-Trp and SD/-Trp/-X-a-gal, and incubated at 30°C for 4 days.

Ethylene treatment

Ethephon (ET), which emits ethylene when dissolved in water, was used as a substitute for ethylene. Leaves from normally grown 3-to 4-week-old plants were used during the trefoil stage. Compared to the leaves sprayed with the water as negative control, Ethylene treatment was performed by spraying the leaves with the mixture of 1 mM/L ethephon (Solarbio Bio, Beijing, China). before leaf tissue was sampled, All the control and treated plants were enclosed in plastic bags for different time and place in a sealed chamber at 25°C with a 16-h-light/8-h-dark photoperiod. The whole plants were harvested 0, 2, 4, 6, 8, 10, 12 and 24h after treatments. immediately frozen in liquid nitrogen and stored frozen at -80°C until use.

Abbreviations

CDS: the coding sequence; ERF: ethylene responsive factor; ET: Ethephon; GFP: Green fluorescent protein; ORF: open reading frame; PCD: programmed cell death; VIGS: virus induced gene silencing

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Competing Interests

The authors declare there are no competing interests.

Funding

This research is financially supported by the National Key R and D Plan of China (No. 2018YFD0100402) and National Natural Science Foundation of China (No. 31621005 and No. 31901581). The funding bodies did not play a role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript, but just provide the financial support.

Author contribution statement

Z.X.L, H.L.C and C.F.W conceived and designed the research; C.F.W performed the experiments and draft the manuscript; S.Y.L helped to prepared figures; D.Y.Z., Y.P.Z and L.M.L provided the discussion and propose. Q.L.W collected and analyzed test data and G.L.S supervised all of the project and corrected the manuscript. All authors read and approved the final manuscript.

Acknowledgments

We are grateful to the State Key Laboratory of Cotton Biology, Institute of Cotton research, Chinese Academy of Agricultural Sciences for providing the seeds of *G. hirsutum*, and the base vectors for assembling the transformation constructs.

References

1. Sunilkumar G, Campbell LM, Puckhaber L, Stipanovic RD, Rathore KS. Engineering cottonseed for use in human nutrition by tissue-specific reduction of toxic gossypol. *Proc. Natl. Acad. Sci. USA*. 2006; 103:18054-18059.
2. Wang K, Wang Z, Li F, Ye W, Wang J, Song, G, et al. **The draft genome of a diploid cotton *Gossypium raimondii***. *Nat. Genet.* 2012; 44:1098-1103.
3. Bao Y, Hu G, Flagel LE, Salmon A, Bezanilla M, Paterson AH, et al. **Parallel up-regulation of the profilin gene family following independent domestication of diploid and allopolyploid cotton (*Gossypium*)**. *Proc. Natl Acad. Sci. USA*. 2011; 108: 21152-21157.
4. Wendel JF. **Genome evolution in polyploids**. *Plant Mol. Biol.* 2000; 42:225-249.
5. Zhang, HB, Li Y, Wang B, Chee PW. **Recent advances in cotton genomics**. *Int. J. Plant Genomics*. 2008; 742304.
6. Yu J, Kohel RJ, Smith CW. **The construction of a tetraploid cotton genome wide comprehensive reference map**. *Genomics*. 2010; 95:230–240.
7. Bell AA, Stipanovic RD. **The chemical composition, biological activity and genetics of pigment glands in cotton**. *Proc. Beltwide Cotton Prod. Res. Conf.* 1977; 244-258.
8. Zhang WJ, Xu ZR, Pan XL, Yan XH, Wang YB. **Advances in gossypol toxicity and processing effects of whole cottonseed in dairy cows feeding**. *Livest. Sci.* 2007; 111: 1-9.
9. Cai YF, Xie YF, Liu JG. **Glandless seed and glanded plant research in cotton**. A review. *Agron Sustain Dev.* 2010; 30: 181-190.

10. Gao, W, Long L, Zhu LF, Xu L, Gao WH, Sun LQ, et al. **Proteomic and virus-induced gene silencing (VIGS) Analyses reveal that gossypol, brassinosteroids, and jasmonic acid contribute to the resistance of cotton to *Verticillium dahliae*.** Mol. Cell Proteomics. 2013; 12:3690-3703.
11. Tian X, Ruan JX, Huang JQ, Yang CQ, Fang X, Chen ZW, et al. **Characterization of gossypol biosynthetic pathway.** Proc. Natl Acad. Sci. USA. 2018; 115: E5410-E5418.
12. McMichael SC. **Glandless boll in upland cotton and its use in the study of natural crossing.** Agronomy Journal. 1954; 46:527-528.
13. McMichael SC. **Hopi Cotton, a Source of Cottonseed Free of Gossypol Pigments 1.** Agronomy journal. 1959; 51: 630-630.
14. McMichael SC. **Combined effects of glandless genes gl2 and gl3 on pigment glands in the cotton plant.** Agronomy Journal. 1960; 52:385-386.
15. Gutierrez M, Vrdoljak J, Ricciardi A. **Development of gossypol-glandless strains of cotton. In Induced Mutations and Plant Improvement,** pp. 1972;397-404. *Vienna, Austria: International Atomic Energy Agency.*
16. Lee JA. **The genomic allocation of the principal foliar-gland loci in *Gossypium hirsutum* and *Gossypium barbadense*.** Evolution.1965; 19:182-188.
17. McCarty JC, Hedin PA, Stipanovic RD. **Cotton *Gossypium* spp. plant gossypol contents of selected *GL₂* and *GL₃* alleles.** *J. Agric. Food Chem.* 1996; 44: 613-616.
18. Scheffler JA, Romano GB. **Modifying gossypol in cotton (*Gossypium hirsutum* L.): a cost effective Method for small seed samples.** J Cotton Sci. 2008; 12: 202-209.
19. Scheffler JA, Romano GB. **Registration of GVS1, GVS2, and GVS3 upland cotton lines with varying gland densities and two near-isogenic lines, GVS4 and GVS5.** J. Plant Reg.2012; 6:190-194.
20. Miravalle RJ, Hyer AH. **Identification of the Gl2 gl2 Gl3gl3 genotype in breeding for glandless cottonseed.** Crop Sci. 1962; 2: 395-397.
21. Endrizi JE, Turcotte EL, Kohel RJ. **Genetics, cytology, and evolution of *Gossypium*.** Adv. Genet. 1985; 23: 271-375.
22. Percy R, Hendon B, Bechere E, Auld D. **Qualitative genetics and utilization of mutants.** In Cotton (Fang, D.D. and Percy, R.G., eds), pp. 2015; 155-186. Madison, WI: American Society of Agronomy, Inc., Crop Science Society of America, Inc., and Soil Science Society of America, Inc.
23. Lee JA. **Genetical studies concerning the distribution of pigment glands in the cotyledons and leaves of upland cotton.** Genetics.1962; 47:131-142
24. Lusas EW, Jividen GM. **Glandless cottonseed: a review of the first 25 years of processing and utilization research.** *J. Am. Oil Chem. Soc.* 1987; 64: 839-854.
25. Barrow JR, Davis DD. **Gl₂^s-a new allele for pigment glands in cotton.** Crop Sci. 1974; 14:325-326.
26. Kohel RJ, Lee JA. **Genetic analysis of Egyptian glandless cotton.** Crop Sci. 1984; 24:1119–1121.
27. Zhang TZ, Zhang XL, Jin L, Chen ZX, Guo WZ. **Genetic identification of a new gland forming gene in upland cotton.** Acta Agron Sin. 2001; 27: 75-79

28. Zhu SJ, Reddy N, Jiang YR, Ji DF. **Breeding, introgression and inheritance of delayed gland morphogenesis trait from *Gossypium bickii* into upland cotton germplasm.** Chin Sci Bull. 2004; 49: 2470-2476
29. Afifi A, Bary A, Kamel S. and Heikal I. **Bahtim 110, a new strain of Egyptian cotton free from gossypol.** Empire Cotton Growing Rev. 1966; 43:112-120.
30. Carvalho LD, Vieira RdM. **Expression of the *Gossypium barbadense* Gl_2^e gene in *Gossypium hirsutum* annual cotton.** Rev Oleaginosas Fibrosas. 2000;4: 39-44
31. Dong C, Ding Y, Guo W, Zhang T. **Fine mapping of the dominant glandless Gene Gl_2^e in Sea-island cotton (*Gossypium barbadense* L.).** Chin Sci Bull. 2007; 52: 3105-3109. doi:10.1007/s11434-007-0468-6.
32. Tang CM, Min LF, Pan JJ, Jin SY. **Genetic analysis for Hai1 strain of glandless cotton (*G. barbadense* L.): interaction between Gl_2^e and *Gl1*.** Cotton Sci. Sin. 1996; 8: 138-140.
33. Cheng HI, Lu CR, Yu JZ, Zou CS, Zhang, YP, Wang QL, et al. **Fine mapping and candidate gene analysis of the dominant glandless gene Gl_2^e in cotton (*Gossypium spp.*).** Theor. Appl. Genet. 2016; 129:1347-1355.
34. Ma D, Hu Y, Yang CQ, Liu BL, Fang L, Wan Q, et al. **Genetic basis for glandular trichome formation in cotton.** Nat. Commun. 2016; 7:10456.
35. Janga, MR, Pandeya D, Campbell LM, Konganti K, Villafuerte ST, Puckhaber L, et al. **Genes regulating gland development in the cotton plant.** Plant Biotechnol. J. 2019; 17: 1142-1153.
36. Gao W, Xu FC, Long Lu, Li Y, Zhang JL, Chong L, et al. **The gland localized CGP1 controls gland pigmentation and gossypol accumulation in cotton.** Plant Biotechnology Journal. 2019;doi:10.1111/PBI.13323.
37. Cai Y, Mo J, Zeng Y, Ren W, Xu Y, Wang S, et al. **Cloning of cDNAs of differentially expressed genes in the development of special pigment gland of cotton by suppression subtractive hybridization.** J. Beijing Univ. Forest. 2003; 25, 6-10.
38. Cai YF, Cai XY, Wang QL, Wang P, Zhang Y, Cai CW, et al. **Genome sequencing of the Australian wild diploid species *Gossypium australe* highlights disease resistance and delayed gland morphogenesis.** Plant Biotechnology Journal. pp. 2019; 1-15.
39. Chang PA, Li B, Ni XM, Xie YF, Cai YF. **Molecular cloning and expression analysis of a RanBP2 zinc finger protein gene in upland cotton (*Gossypium hirsutum* L.).** Colloids and Surfaces B: Biointerfaces. 2007; 55: 153-158.
40. Müller M, Munné-Bosch S. **Ethylene Response Factors: A Key Regulatory Hub in Hormone and Stress Signaling.** Plant Physiol [Internet]. 2015; 169: 32–41. Available from: <http://www.plantphysiol.org/lookup/doi/10.1104/pp.15.00677>.
41. Dey S, Corina VA. **Ethylene responsive factors in the orchestration of stress responses in monocotyledonous plants.** Front Plant Sci [Internet]. 2015; 6:1-7.

42. Huang PY, Catinot J, Zimmerli L. **Ethylene response factors in Arabidopsis immunity.** J Exp Bot. 2016; 67:1231–41.
43. Rotenberg D, Thompson TS, German TL, Willis DK. **Methods for effective real-time RT-PCR analysis of virus-induced gene silencing.** J Virol Methods. 2006; 138: 49-59.
44. Xiong L, Schumaker KS, Zhu JK. **Cell signaling during cold, drought, and salt stress.** Plant Cell (Suppl). 2002; 14:165-183.
45. Shinshi H, Usami S, Ohme-Takagi M. **Identification of an ethylene-responsive region in the promoter of a tobacco class I *chitinase* gene.** Plant Mol Biol. 1995; 27:923-932
46. Sessa G, Meller Y, Fluhr R. **A GCC element and a G-box motif participate in ethylene-induced expression of the PRB-1b gene.** Plant Mol Biol. 1995; 28:145-153.
47. Jiang LL, Lian XF, Fan MS. **Role of programmed cell death in adaptation of plant to environmental stress.** Chinese Bulletin of Life Sciences. 2005; 3: 267-270.
48. Liu WZ, Zhou YF, Wang X, Jiao ZJ. **Programmed cell death during pigment gland formation in *Gossypium hirsutum* leaves.** Plant Biology. 2010; 12: 895-902
49. Ni, XL. **Aerenchyma formation and their relationship to programmed cell death in three plants.** (Ph D dissertation). Xibei University. 2014
50. Yuan YL, Chen YH, Tang CM, Jing SR, Liu SL, Pan JJ, et al. **Effects of the dominant glandless gene *Gl₂^e* on agronomic and fiber characters of Upland cotton.** Plant Breed. 2000; 119: 59-64.
51. Livak KJ, Schmittgen TD. **Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta CT}$ Method.** Methods. 2001; 25: 402-408. doi:10.1006/meth.2001.1262. PubMed: 11846609.
52. Burch-Smith TM, Anderson JC, Martin GB, Dinesh-Kumar SP. **Applications and advantages of virus-induced silencing for gene function studies in plant.** Plant J. 2004; 39: 734-746.
53. Gao X, Wheeler T, Li Z, Kenerley CM, He P, Shan L. **Silencing GhNDR1 and GhMKK2 compromises cotton resistance to Verticillium wilt.** Plant J. 2011; 66:293–305
54. Zhang JX, Wang FR, Zhang CY, Zhang JH, Chen Y, Liu GD, et al. **A novel VIGS method by agro inoculation of cotton seeds and application for elucidating functions of GhBI-1 in salt-stress response.** Plant Cell Reports. 2018; 37: 1091-1100.
55. Gao XQ, Shan LB. **Functional Genomic Analysis of Cotton Genes with Agrobacterium-Mediated Virus-Induced Gene Silencing.** Methods Mol Biol. 2013; 975:157-165. doi:10.1007/978-1-62703-278-0_12.
56. Yi CX, Zhang J, Chan KM, Liu XK, Hong Y. **Quantitative real-time PCR assay to detect transgene copy number in cotton (*Gossypium hirsutum*).** Anal. Biochem. 2008; 375:150-152.
57. Stefan O, Grit K, Jonathan G. **Increased Terpenoid Accumulation in Cotton (*Gossypium hirsutum*) Foliage is a General Wound Response.** J Chem Ecol. 2008; 34:508-522.

Figures

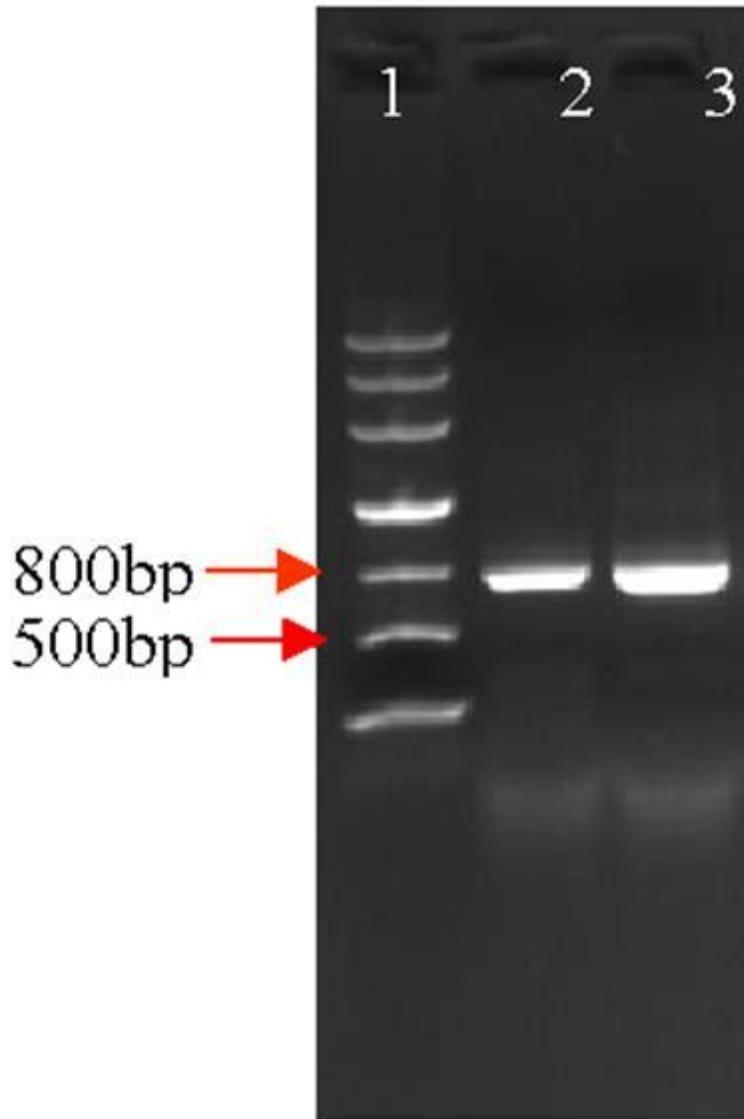


Figure 1

Amplification of the full-length cDNA of GhERF105. 1: DNA marker; 2-3: the full-length cDNA of GhERF105

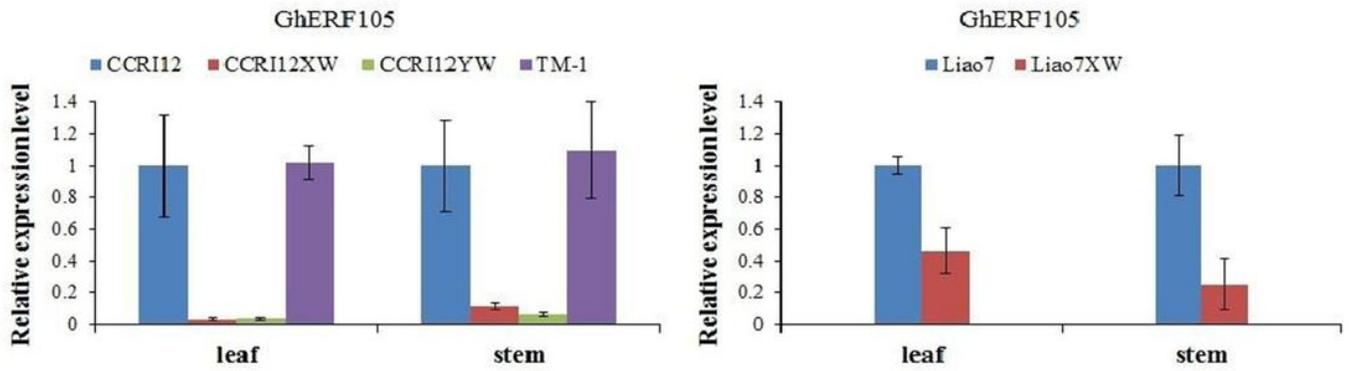


Figure 2

Real-time RT-PCR expression analysis of GhERF105 in the leaf and stem of CCR112, CCR112XW, CCR112YW, L7, L7XW and TM-1. Error bars represent the standard of the mean values of three biological replicates.

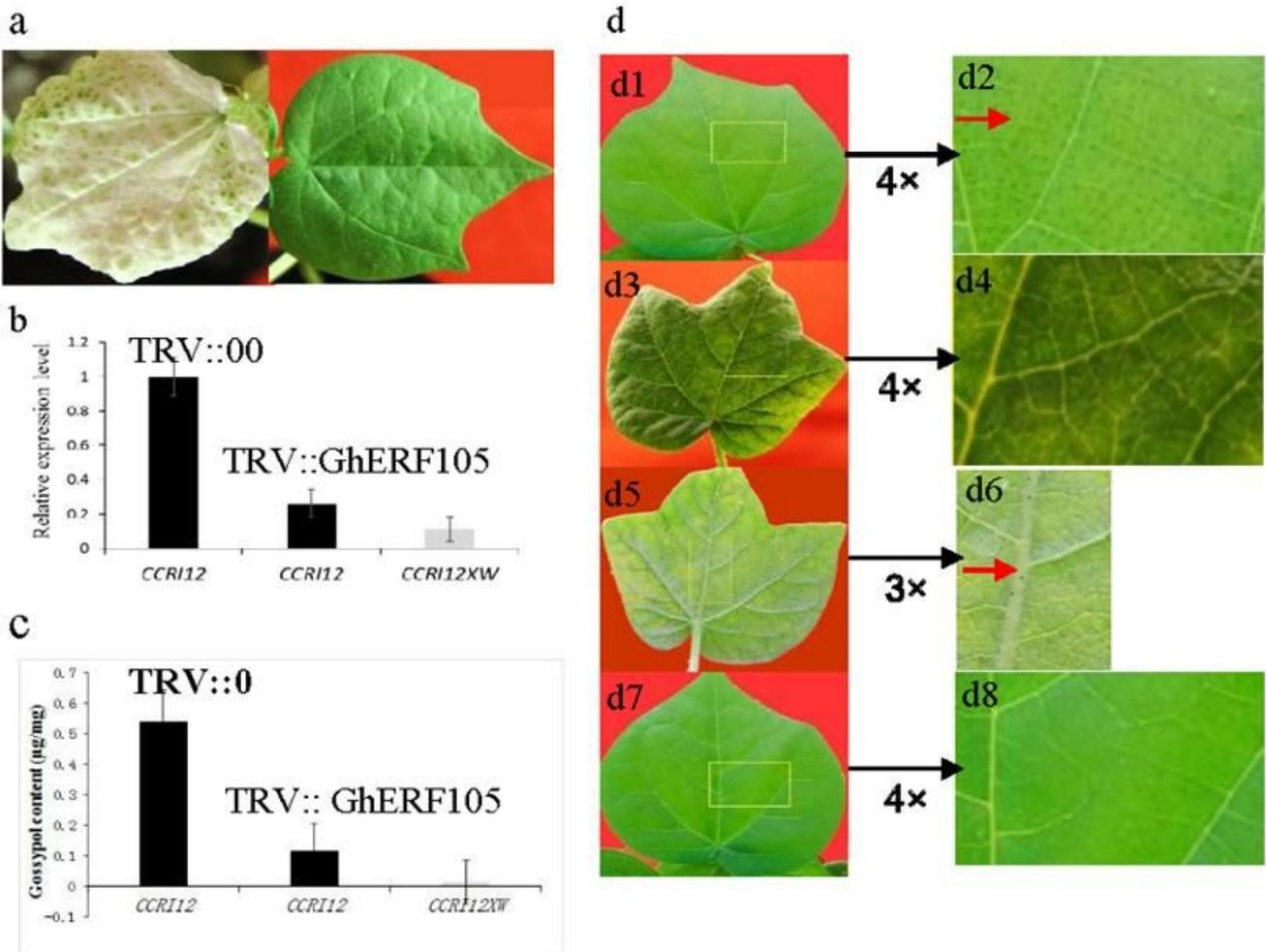


Figure 3

Functional characterization of GhERF105 by VIGS. a The photo-bleaching phenotypes of cotton seedlings in CCRI12 inoculated with pTRV::GhPDS and empty vectors. TRV::GhPDS and TRV::00 are the positive control and negative control, respectively. b Relative transcript levels of GhERF105 of leaf inoculated with pTRV::GhERF105 or empty vector control. c The gossypol content in empty vector (TRV::00) and in the GhERF105-silenced leaves of CCRI12.and CCRI12XW, Actin was used as an internal control. d d1-d2 Phenotypes of *Gossypium hirsutum* CCRI12 inoculated with pTRV::00 vector. d3-d6 Phenotypes of *Gossypium hirsutum* CCRI12 inoculated with pTRV:: GhERF105 vector. d7-d8 Phenotypes of *Gossypium hirsutum* CCRI12XW. d1-d8 are enlarged versions of the positions indicated by the yellow box in Fig.3. d1 correspond to d2, d3 correspond to d4, d5 correspond to d6 and d7 correspond to d8. the red arrow indicates the location of the glands on the leaf. Scale bars, 1mm. Each bar value represents mean \pm SD of three independent experiments. Error bars are the SD of three biological repeats.

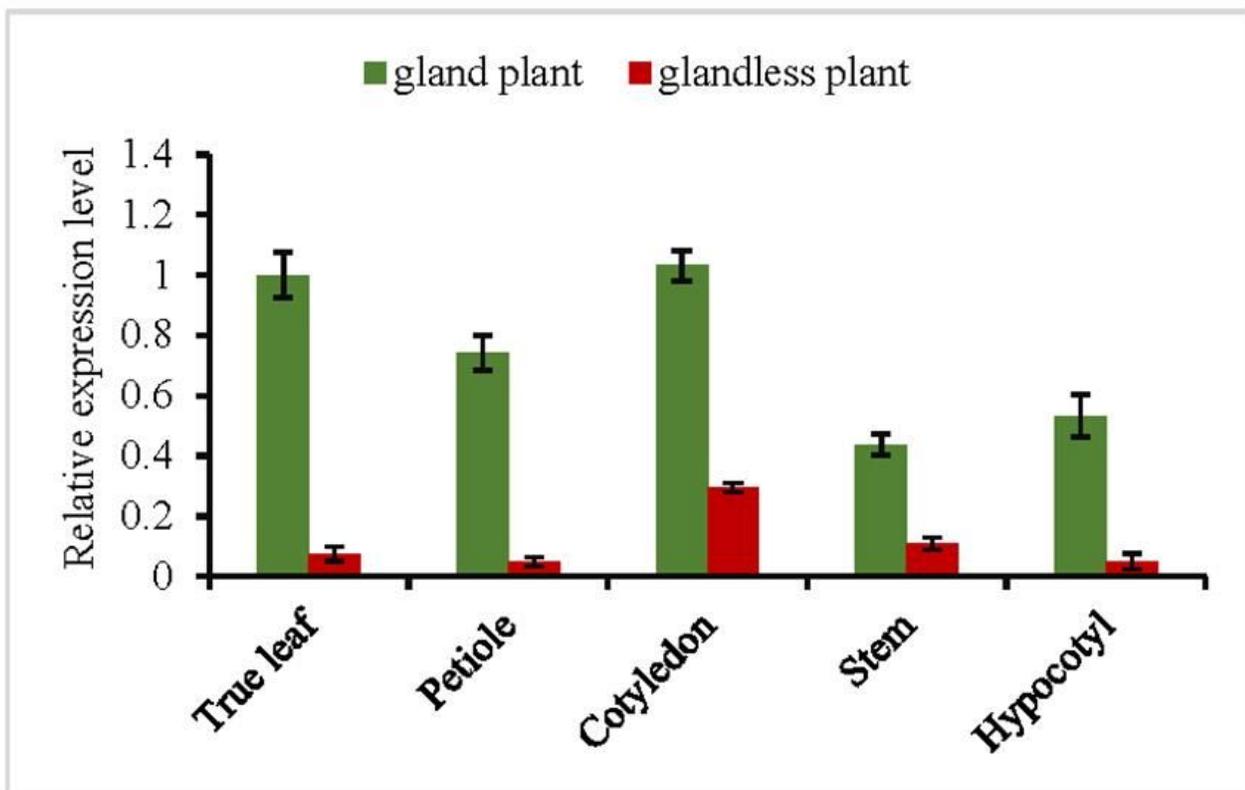


Figure 4

Spatiotemporal expression analysis of GhERF105 gene in different tissue of glanded plant and glandless plant.

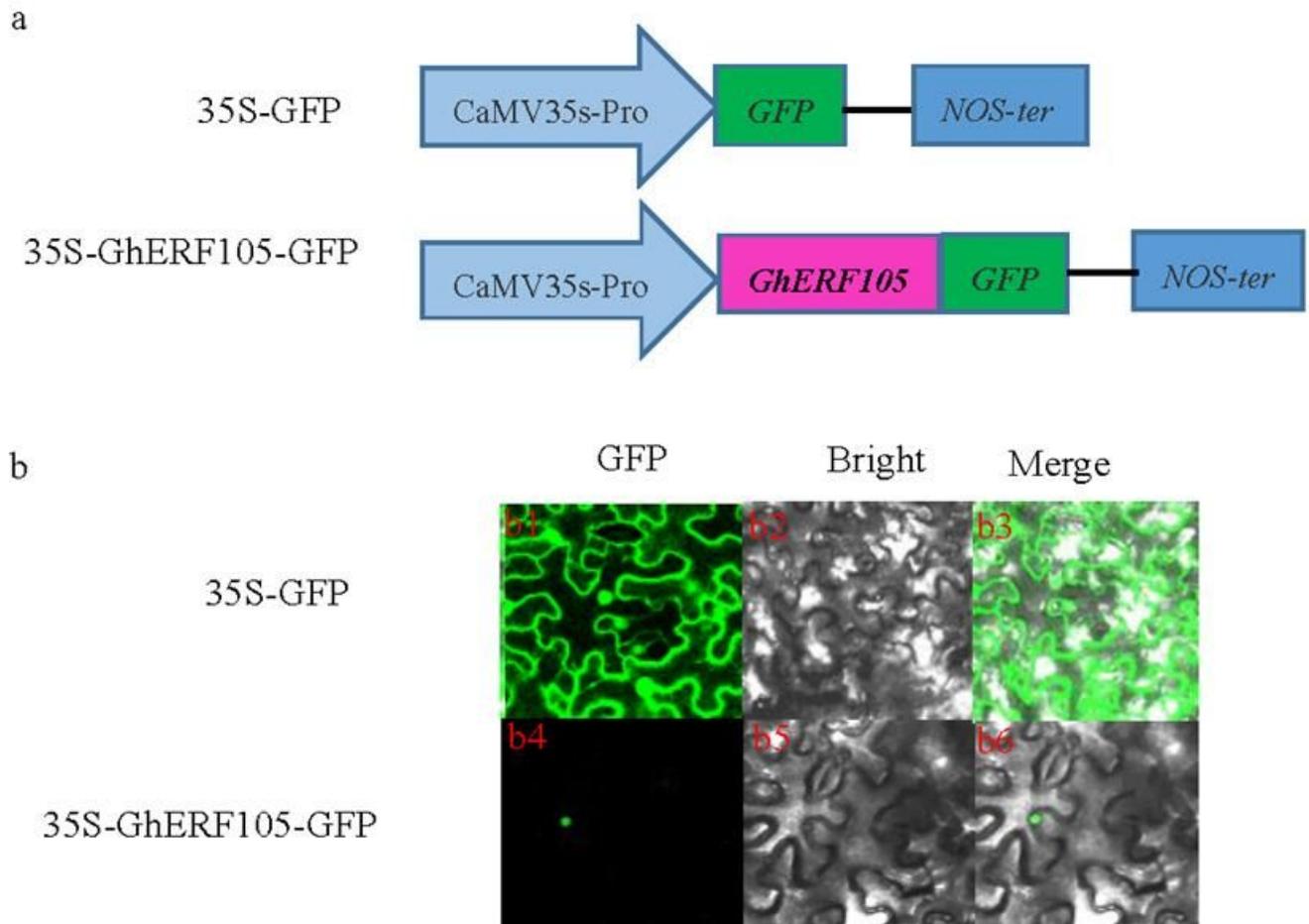


Figure 5

Subcellular localization of GhERF105. a The pBI121::GhERF105 fusion expression vector was constructed. b Subcellular localization of GFP (b1, b2 and b3) and GhERF105::GFP fusion protein (b4, b5 and b6) in tobacco cells Bar=100µm

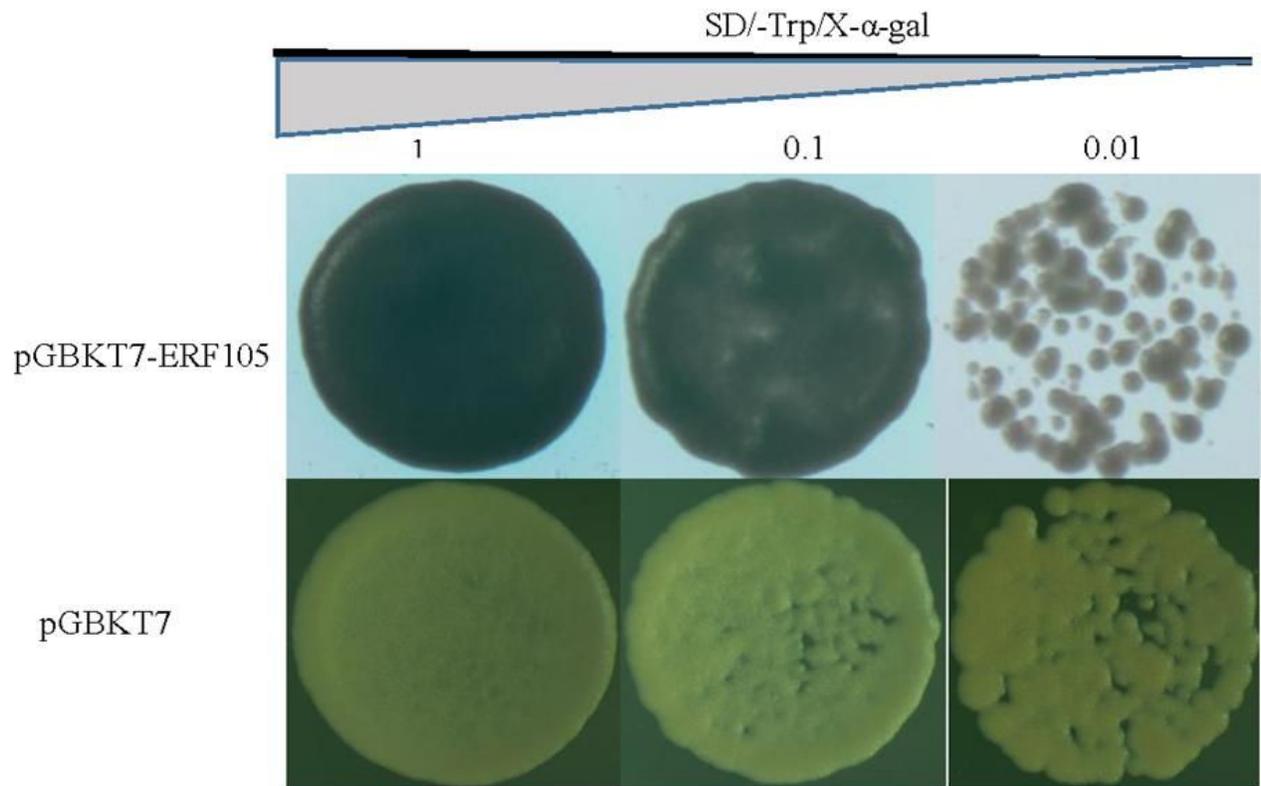


Figure 6

Transcription activation of GhERF105 in yeasts. The constructs pGBKT7 and pGBKT7-GhERF105 were transformed in yeasts and grew for 3-4days on the selected medium as indicated, Bar=200 μ m

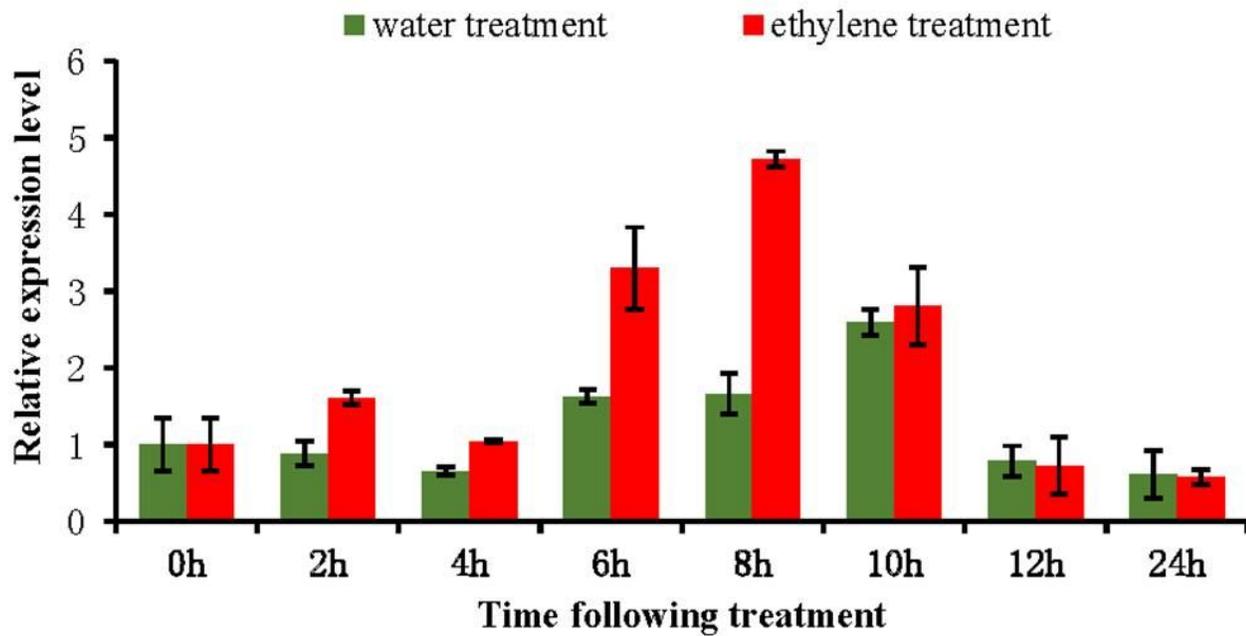


Figure 7

The expression of GhERF105 gene in leave of glanded plant under ethylene treatment.

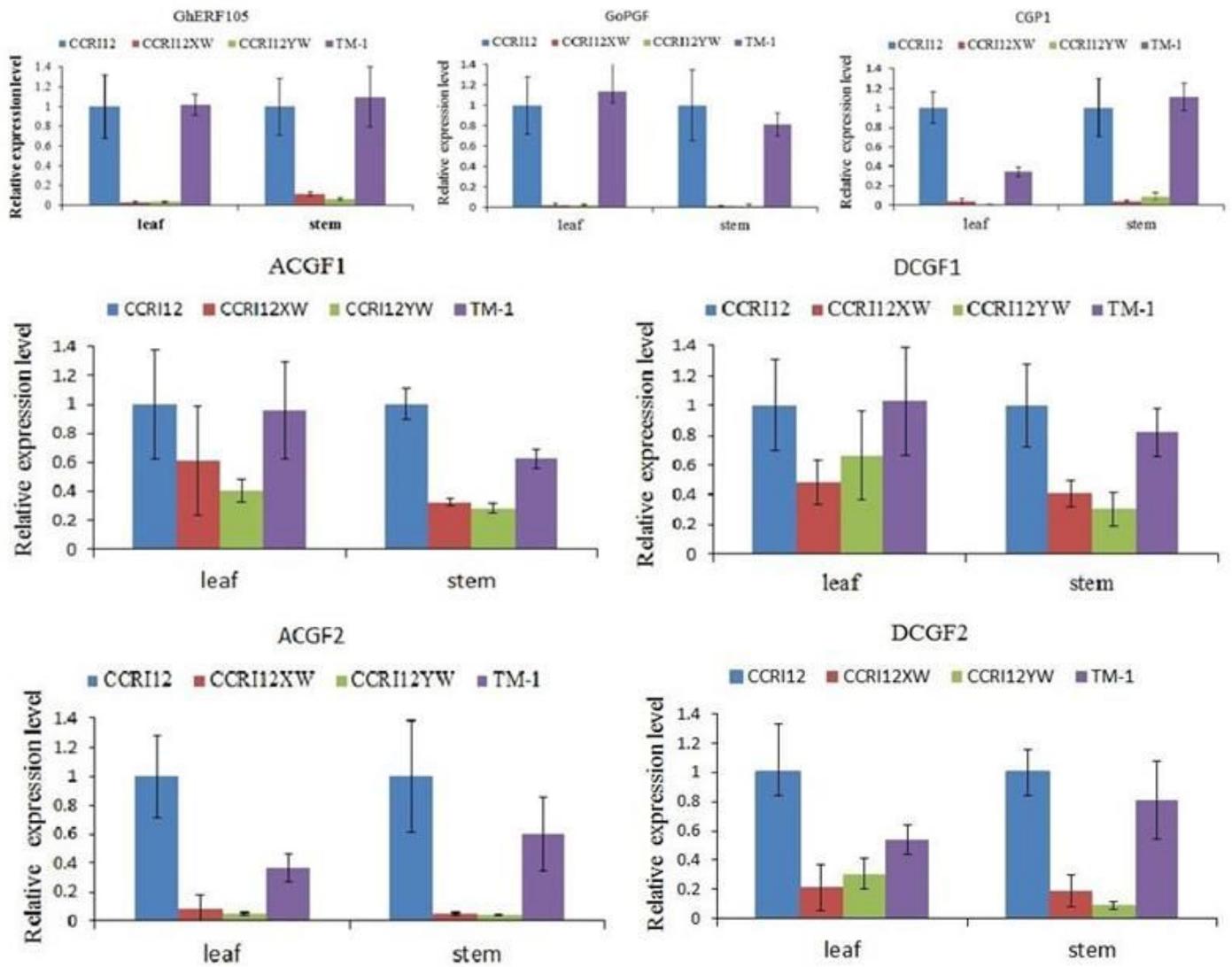


Figure 8

Expression analysis of GhERF105, GoPGF, CGF1, CGF2 and CGP1 genes in the leaf and stem of glanded plant and glandless plant.

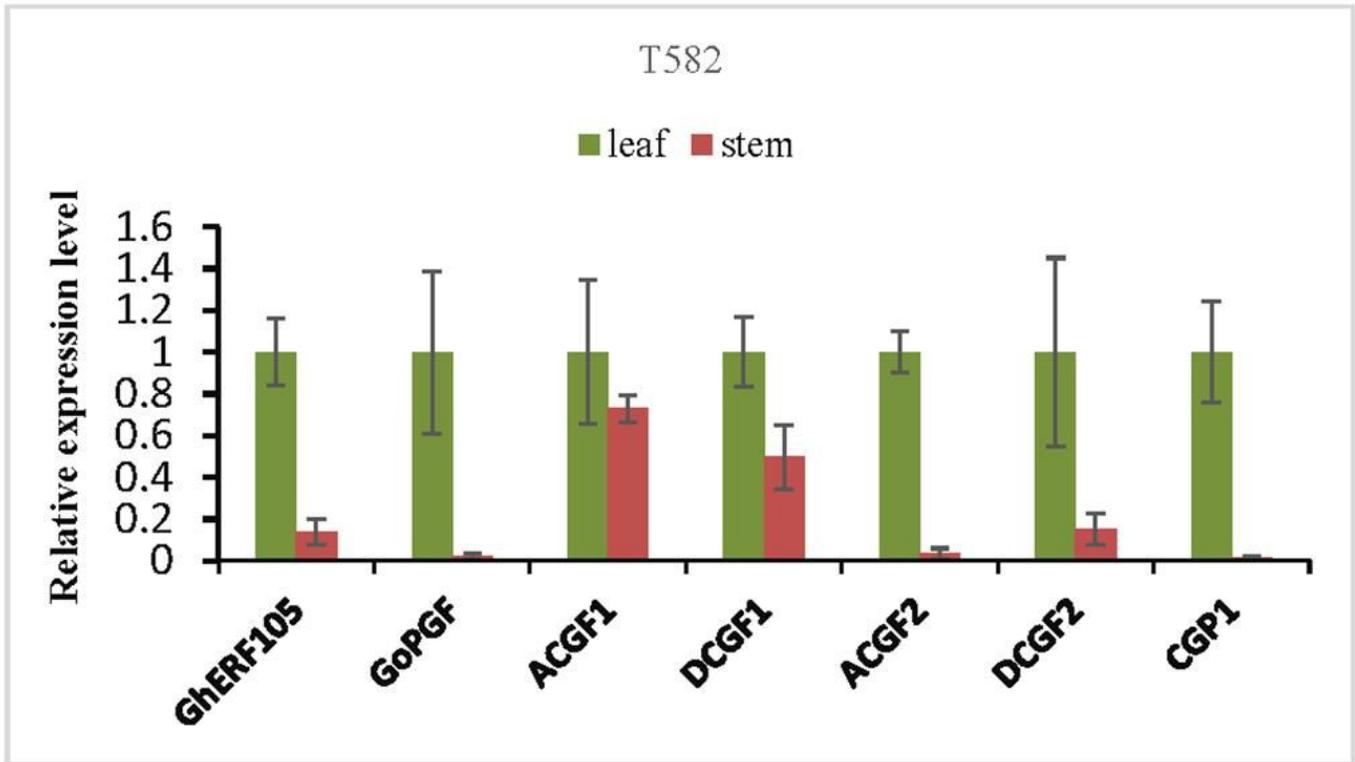


Figure 9

Expression analysis of GhERF105, GoPGF, CGF1, CGF2 and CGP1 genes in the leave and stem of T582 cotton cultivar.

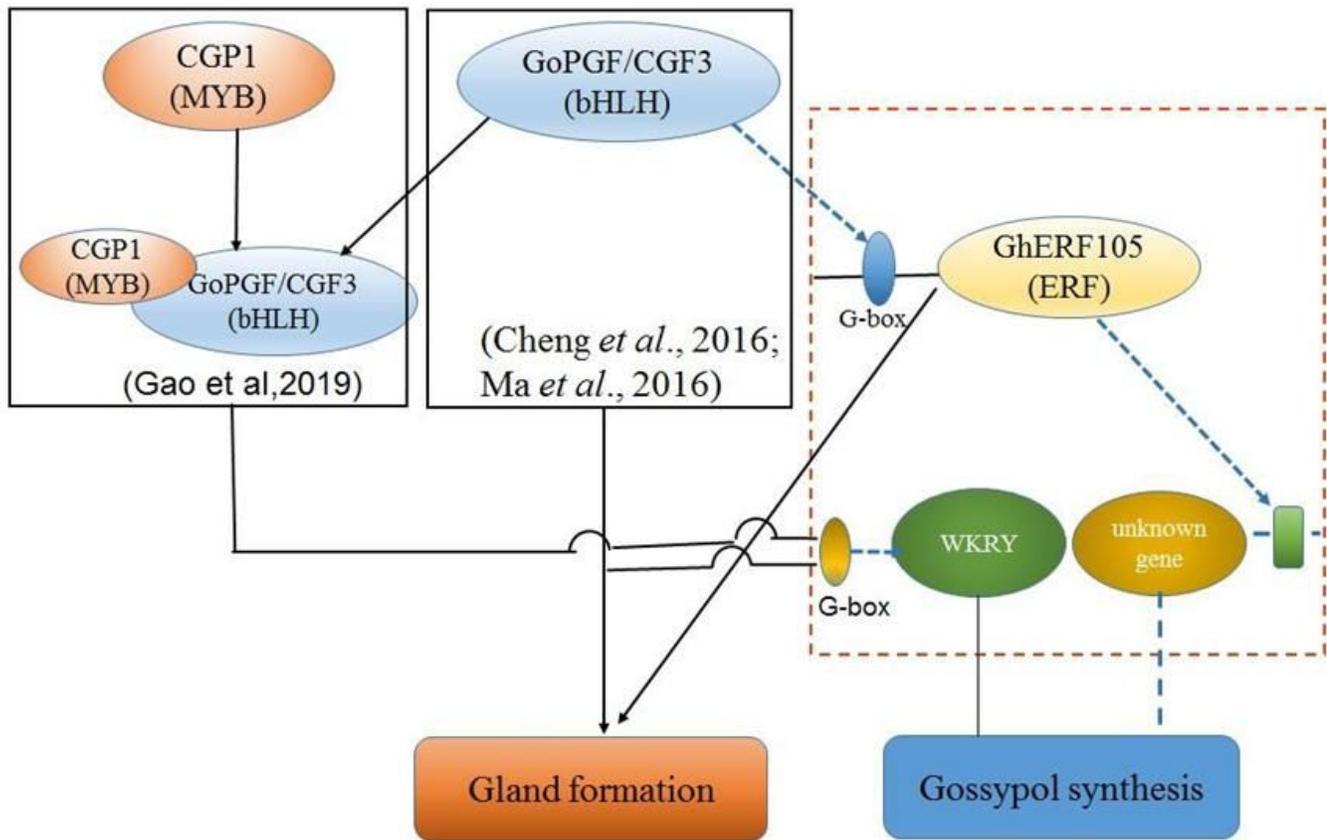


Figure 10

Schematic model illustrating the proposed functions of GhERF105 and in cotton gland pigmentation formation.

Supplementary Files

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

- [supplementaryfigure2.pdf](#)
- [supplementarytables.xls](#)
- [supplementaryfigure1.pdf](#)