GENETICS

The genes determining synthesis of pigments in cotton

Aleksandra Mikhailova[1,2], Ksenia Strygina[2,3], and Elena Khlestkina[2,3]

[1]Department of Genetics and Biotechnology, Faculty of Biology, Saint Petersburg State University, Universitetskaya nab., 7–9, Saint Petersburg, 199034, Russian Federation
[2]Federal Research Center N. I. Vavilov All-Russian Institute of Plant Genetic Resources, B. Morskaya Str., 42–44, Saint Petersburg, 190000, Russian Federation
[3]Federal Research Center Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences, Pr. Akademika Lavrentyeva, 10, Novosibirsk, 630090, Russian Federation

Address correspondence and requests for materials to Aleksandra Mikhailova, a.mikhailova@vir.nw.ru

Abstract

Naturally coloured cotton is environmentally friendly, since bleaching and chemical dyeing are not needed during textile production. Studying molecular-genetic mechanisms underpinning pigment production may facilitate breeding cotton with coloured fibre. In the current review we summarize the known data on structural and regulatory genes involved in biosynthesis of flavonoid pigments proanthocyanidins (PAs) in brown and caffeic acid (CA) derivates in green fibre. The first chapter considers the first studies on fibre cotton inheritance, from the beginning of the last century. Then, we briefly review the biochemical and physico-chemical methods proving the presence of PAs in brown fibre and derivatives of CA in green cotton fibre. The biochemical analysis of coloured cotton fibre was followed by genetic studies of structural genes coding for enzymes participating in PAs and CA biosynthesis, transport and oxidation processes. We summarize the data on the genes coding for transcription factors from the MBW (MYB-bHLH-WD40) regulatory complex, which controls flavonoid biosynthesis in coloured cotton fibre. The regulatory gene most interesting as a target for markers-assisted breeding and genome editing is *GhTT2-3A.*

Keywords: brown fibre, caffeic acid, flavonoids, green fibre, *Gossypium*, MBW regulatory complex, proanthocyanidins

Introduction

Cotton (*Gossypium* L.) is an important crop. The raw materials obtained from cotton are used in the textile and military industries due to the unique properties of cotton fibres, as well as in the pharmaceutical, food, and chemical industries due to the high content of some valuable metabolites. (Gong et al., 2018). Vegetable oils isolated from cotton seeds contain a significant amount of such fatty acids as palmitic, stearic, arachidic, oleic and linoleic acid (Dinesh K. Agarwal et al., 2003). The biological activity of gossypol extracted from cotton root, bolls and floral organs leads to it being used as a valuable pharmaceuticals component (Wang et al., 2009). Cotton fibre serves as an indispensable source of raw materials in the fabric manufacture. However, bleaching and chemical dyeing of fabrics during production contribute to environment pollution. In this regard, naturally coloured cotton fibre which is also called "eco-friendly cotton" can be used for the production of environmentally friendly fabrics (Nimon and Beghin, 1999). Furthermore, the use of naturally coloured cotton reduces the cost of fabric, since expensive bleaching and chemical dyeing are no longer needed.

Citation: Mikhailova, A., Strygina, K., and Khlestkina, E. 2019. The genes determining synthesis of pigments in cotton. Bio. Comm. 64(2): 133–145. https://doi.org/10.21638/spbu03.2019.205

Author's information: Aleksandra Mikhailova, Junior Researcher, orcid.org/0000-0003-4565-1539; Ksenia Strygina, PhD, Senior Researcher, orcid.org/0000-0001-6938-1348; Elena Khlestkina, Dr. of Sci. in Biology, Professor RAS, Head of Laboratory, orcid.org/0000-0002-8470-8254

Manuscript Editor: Anton Nizhnikov, Department of Genetics and Biotechnology, Faculty of Biology, Saint Petersburg State University, Saint Petersburg, Russia

Received: May 31, 2019;

Revised: July 10, 2019;

Accepted: July 17, 2019;

Funding: This work was conducted within the framework of the state task regarding Theme № 0481-2019-0001 "Genomic and postgenomic technologies for searching for new genetic markers of economically important traits and new allelic variants of agricultural valuable genes in the gene pool of cultivated plants and their wild relatives".

Competing interests: The authors have declared that no competing interests exist.

Fig. 1. The examples of white and brown cotton samples from Herbarium collection of N. I. Vavilov All-Russian Institute of Plant Genetic Resources.

The current review summarizes known data on molecular-genetic mechanisms underpinning pigmentation of fibre during cotton plant development. The metabolic pathways as well as structural and regulatory genes related to pigment synthesis in green and brown cotton fibre are considered.

1. Green and brown fibre colour — the first steps in the investigation

Coloured cotton fibres are characterised by different shades from cream to rust, but green and brown fibres are prevalent (Dutt et al., 2004). Derivatives of caffeic acid accumulating in the suberin layer are known to provide green fibre colour (Yatsu, Espelie, and Kolattukudy, 1983; Schmutz et al., 1993; Schmutz, Jenny, and Ryser, 1994; Feng et al., 2017). Brown pigmentation of cotton fibre is caused by the presence of proanthocyanidins (PAs) in cell vacuoles (Xiao et al., 2007, 2014; Feng et al., 2014; Malik et al., 2015) (Fig. 1).

The first investigations related to the genetic aspects of fibre colour date from the beginning of the last century. It was established that the green fibre trait is dominant over the brown and white colour traits (Harland, 1932; Hutchinson and Silow, 1939). Later, it was shown that the green colour fibre trait is regulated by one gene with incomplete dominance (Richmond, 1943).

At the same time, after crossing uncoloured and reddish-brown *Gossypium* plants it was observed that the genotypes with cream and light-brown fibres are heterozygous due to the splitting ratio and incomplete dominance (Balls, 1908; Kottur, 1923; Richmond, 1943;

Zhang et al., 2002). After further research it was shown that the formation of brown cotton fibre is controlled by a single locus containing genes with incomplete dominance (Ware, 1932).

Twelve years later, it was reported that there are six *Lc* genes (*Lc1–Lc6*) that determine fibre colour in the *Gossypium* genome (Silow, 1944). Further monosomic analysis showed that the *Lc1* gene, which is responsible for brown fibre colour, is located in chromosome 7 (Endrizzi and Taylor, 1968).

In 1985 it was first established that condensed tannins or proanthocyanidins (PAs) are responsible for brown pigmentation in cotton fibre (Ryser and Holloway, 1985). In 1994 caffeic acid (CA) derivatives were extracted from green cotton fibre for the first time (Schmutz et al., 1994). After this discovery, the study of the molecular mechanisms controlling the biosynthesis of these pigments began.

2. Proanthocyanidins in brown cotton fibre

Flavonoids are plant secondary metabolites derived from the general phenylpropanoids pathway (Schijlen, Ric de Vos, van Tunen, and Bovy, 2004; Koes, Verweij, and Quattrocchio, 2005; Panche, Diwan, and Chandra, 2016). Flavonoids have a broad spectrum of functions in plants including protection against adverse environmental factors such as extremal temperature, ultraviolet radiation damage and pathogens (Alonso-Amelot, Oliveros, and Calcagno-Pisarelli, 2004; Takahashi et al., 2010).

PAs (also condensed tannins) are a class of oligomeric flavonoids with the diphenylpropane common chemical structure (C6-C3-C6), which includes two aromatic rings with a three-carbon bridge that forms a heterocyclic ring (Fig. 2) (Feng et al., 2014; Ma et al., 2016b). PAs have antioxidant and anti-inflammation activities (Salunkhe et al., 1983). In plants, PAs are also responsible for pigmentation of plant tissues and usually accumulate in seed coats, leaves, stems and roots (Prasad, 2000; Peng et al., 2012).

Evidence that PAs are contained in brown cotton fibres was first obtained with the qualitative colour reaction with DMACA (dimethylaminocinnamaldehyde), which is commonly used to detect PAs in plant tissues, leading to blue staining (Harland, 1932; Hutchinson and Silow, 1939; Xiao et al., 2007). The comparative treatment of white and brown cotton fibres with DMACA has demonstrated that brown cotton fibre turned into dark-blue, while the mature white fibre did not show a significant difference in colouration (Li et al., 2013). In addition, the more days after anthesis (DPA) have passed, the darker the brown colour shade of cotton fibre. This observation indicates that the oxidized PAs content increases proportionally with maturation of cotton bolls (Feng et al., 2014).

The structure of PAs units in brown cotton fibre was confirmed using a matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF MS), nuclear magnetic resonance (NMR) and liquid chromatography-mass spectrometry (LC-MS) analysis (Nechepurenko et al., 2009; L. Wang et al., 2016). The results revealed that an initiating unit derived from gallocatechin; all extension units contain 3 hydroxyl groups in B-ring (Bogs, 2005; He, Pan, Shi, and Duan, 2008; Ershik and Buzuk, 2009; Feng et al., 2014; Xiao et al., 2014). The structures of PAs are different in white and brown cotton fibres. The content of prodelphinidin (PD) and proanthocyanidin (PC) are equal in white cotton, while in brown 90 % PD and 10 % PC was found (He et al., 2008; Feng et al., 2014). Moreover, it was detected that the structure of PAs in white cotton was modified by a galloyl group (Xiao et al., 2014; Feng et al., 2014). Also, the MALDI-TOF MS proteomic analysis of brown cotton fibre identified that 21 proteins were related to secondary metabolism processes and 15 of them were associated with the flavonoid biosynthesis pathway (Li et al., 2013).

PAs BIOSYNTHESIS DURING THE PHENYLPROPANOID AND FLAVONOID PATHWAY

Structural genes involved in PAs biosynthesis. PAs are synthesized during the flavonoid pathway in different plant tissues (Debeaujon, 2003; Bogs, 2005; He et al., 2008). Identified genes involved in this pathway in *G. hirsutum* are listed in Table 1.

Phenylalanine is a precursor for all flavonoids, including PAs (Peng et al., 2012; Li et al., 2013; Thomas and ElSohly, 2016). The phenylpropanoid pathway begins with transformation of phenylalanine by phenylalanine ammonia-lyase (PAL), cinnamate 4-hydroxylase (C4H) and 4-coumarate CoA ligase (4CL) into 4-Coumaroyl-CoA (Fig. 2). 4-Coumaroyl-CoA exposed such enzymes as chalcone synthase (CHS), chalcone isomerase (CHI), flavanone hydroxylase (F3H), flavonoid 3'-hydroxylase (F3'H), flavonoid 3', 5'-hydroxylase (F3'5'H) and dihydroflavonol 4-reductase (DFR) to leucoanthocyanidins (Fig. 2). Leucoanthocyanidins and anthocyanidins are converted into flavan-3-ols by leucoanthocyanidin reductase (LAR) and anthocyanin reductase (ANR), respectively (Fig. 2). Anthocyanidin synthase (ANS) and anthocyanidin reductase (ANR) lead to anthocyanidin formation from leucoanthocyanidin — a PAs precursor (Bogs, 2005; Peng et al., 2012; Jiafu Tan et al., 2013; Malik et al., 2015). PAs represent the polymer based on flavan-3-ol units (Nechepurenko et al., 2009; Feng et al., 2014; Xiao et al., 2014).

The flavonoid biosynthesis pathway determines pigmentation in cotton fibre (Feng et al., 2013, 2014; Liu et al., 2018). Recently it was demonstrated that overexpres-

Table 1. The structural genes which are involved in PAs biosynthesis

G. hirsutum gene	Chromosome; Gene ID (CottonFGD*)	The protein family	Function	References
GhPAL	A01; Gh_A01G1839 D01; Gh_D01G2080	PAL/histidase family	Phenylalanine ammonia-lyase activity	Feng et al., 2017
GhPAL2	A06; Gh_A06G0667 D06; Gh_D06G0758			Qin et al., 2017
GhC4H1	A10; Gh_A10G1590 D10; Gh_D10G1845	Cytochrome P450	Oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen; iron-ion binding; heme binding	Ling Fan and Wen-Ran Hu, 2012
GhC4H2	A13; Gh_A13G2057 D13; Gh_D13G2458			Ling Fan and Wen-Ran Hu, 2012; Feng et al., 2013
GhCHS	A10; Gh_A10G1079 D10; Gh_D10G1429	Chalcone/stilbene synthases family	Transferase activity (transferring acyl groups other than amino-acyl groups)	Feng et al., 2013
GhCHI	A05; Gh_A05G3491 D04; Gh_D04G1926	Chalcone-flavonone isomerase family protein	Catalyses the intramolecular cyclization of bicyclic chalcones into tricyclic (S)-flavanones	Xiao et al., 2007; Li et al., 2013; Malik et al., 2015
GhF3H	A12; Gh_A12G0549 D12; Gh_D12G0566	2-oxoglutarate-dependent dioxygenases superfamily	Oxidoreductase activity	Xiao et al., 2007; Malik et al., 2015
GhF3'H	A12; Gh_A12G2650 D12; Gh_D12G1798	Cytochrome P450	Oxidoreductase activity	Feng et al., 2013
GhF3'5'H	A07; Gh_A07G1098 D07; Gh_D07G1197			
GhDFR	A05; Gh_A05G1647 D05; Gh_D05G1836	Reductase–epimerase–hydrogenase	Catalytic activity; coenzyme binding	Xiao et al., 2007; Malik et al., 2015
GhANS	A08; Gh_A08G1593 D08; Gh_D08G1902	2-oxoglutarate-dependent dioxygenases superfamily	Oxidoreductase activity	
Gh3GT	A05; Gh_A05G3537 D04; Gh_D04G0070	Glycosyltransferase family	Catalyzes the glycosylation of flavonoids at the 3-O-position	Liu et al., 2018
GhANR	A05; Gh_A05G1424 D05; Gh_D05G1596	Reductase–epimerase–hydrogenase	Catalytic activity; coenzyme binding	Xiao et al., 2007; Malik et al., 2015; Yan et al., 2018
GhLAR	A12; Gh_A12G1558 D12; Gh_D12G1686	NAD(P)-binding domain superfamily	Oxidoreductase activity	Yan et al., 2018
Gh4CL1	A02; Gh_A02G1344 D03; Gh_D03G0479	AMP-binding	Ligase activity	Feng et al., 2017; Ling Fan and Wen-Ran Hu, 2012
Gh4CL2	A02; Gh_A02G1344 D03; Gh_D03G0479			
GhCAD1	A02; Gh_A02G1320 D03; Gh_D03G0457	Zinc-containing alcohol dehydrogenase family	Cinnamyl-alcohol dehydrogenase activity; zinc-ion binding	Ling Fan and Wen-Ran Hu, 2012
GhCAD3	A05; Gh_A05G1579 D05; Gh_D05G1757			
GhCAD7	A11; Gh_A11G2005 D11; Gh_D11G1980	Zinc-containing alcohol dehydrogenase family	Cinnamyl-alcohol dehydrogenase activity; zinc-ion binding	Qin et al., 2017; Ling Fan and Wen-Ran Hu, 2012
GhCAD5	A12; Gh_A12G2203 D12; Gh_D12G2382			
GhCAD6	A12; Gh_A12G0193 D12; Gh_A12G0195			

* CottonFGD — Cotton Functional Genomics Database

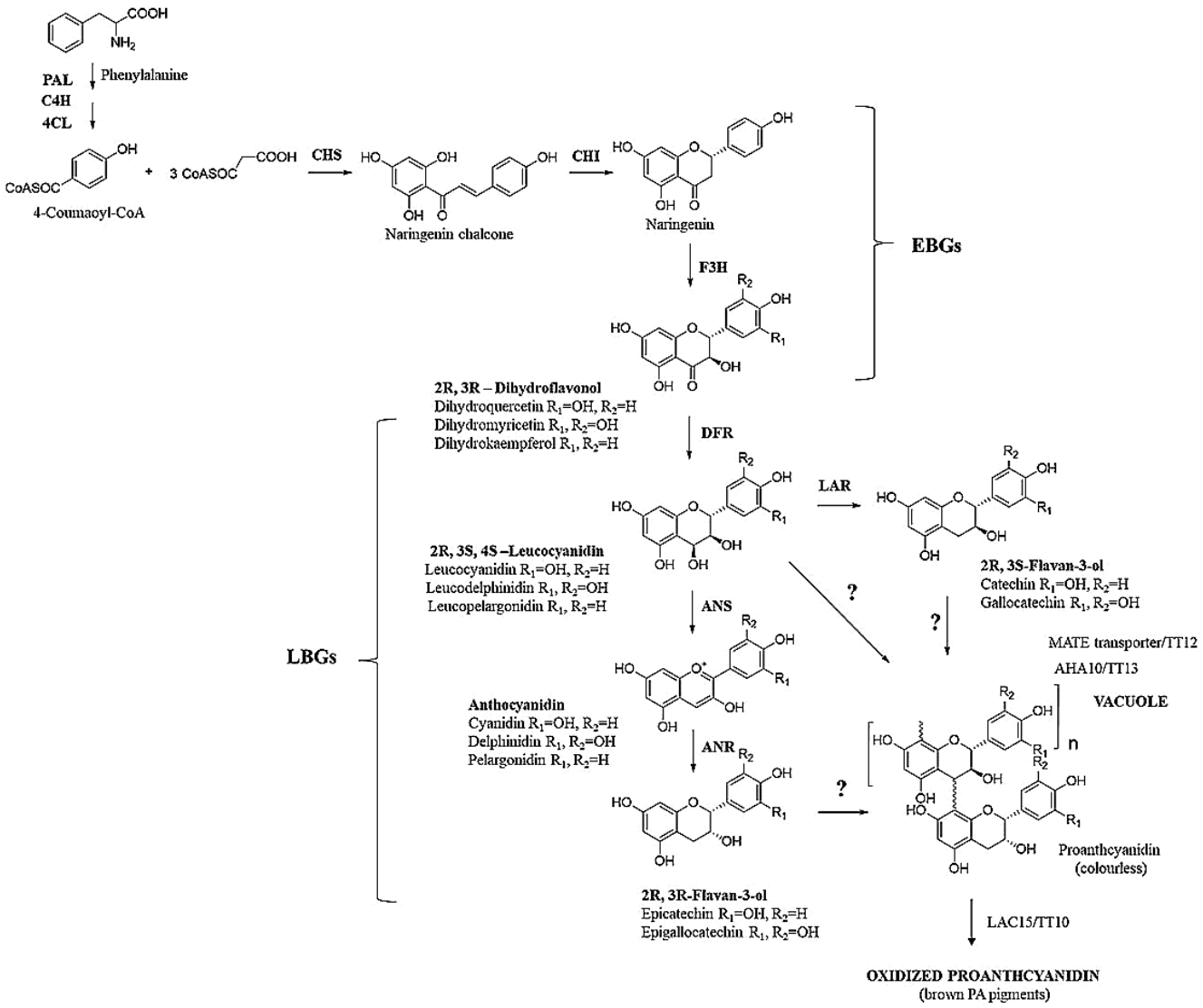

Fig. 2. Anthocyanins and PAs biosynthesis during flavonoid pathway (the scheme are modified from Schijlen, et al. 2004). Enzymes are involved in biosynthesis, abbreviated as: PAL, phenylalanine ammonia-lyase; C4H, cinnamate 4-hydroxylase; 4CL, 4-(hydroxy)cinnamoyl CoA ligase; CHS, chalcone synthase; CHI, chalcone isomerase; F3H, flavanone hydroxylase; DFR, dihydroflavonol 4-reductase; ANS, anthocyanidin synthase; LAR, leucoanthocyanidin reductase; ANR, anthocyanidin reductase. MATE/TT12 multidrug and toxin extrusion family, AHA10/TT13 — H^+-ATPase isoform 10, LAC15/TT10 — laccase 15. MATE and AHA10 participate in transportation of PAs precursors from cytoplasm to vacuole. Early biosynthetic genes and later biosynthetic genes are designate as EBGs and LBGs, respectively. LBs regulated by MBW complex (TT2-TT8-TTG1) in monocots plants, where MBW is Myb-bHLH-WD40 repeat complex.

sion of *Gh3GT* coding the flavonoid 3-glucosyltransferase leads to green cotton fibre formation in brown cotton line (H.-F. Liu et al., 2018). Furthermore, three fibre phenotypes, including brown, white and green colour traits, were observed under downregulation of the *GhCHI* gene in brown cotton using the RNAi-mediated inhibition approach (Abdurakhmonov et al., 2016).

The suppression of the *F3H* gene leads to accumulation of naringenin (NAR), which inhibits the rate of fibre development (Tan et al., 2013). Transformation of F3H-RNA interference segment into the DNA of brown cotton fibre of the *G. hirsutum* line has demonstrated PA content reduction and a fibre shortening compared with control samples (Gong et al., 2014). On the other hand, the overexpression of the *F3H* gene has no effect on cotton fibre length. This experiment has shown that flavonoid metabolism could be associated with fibre pigmentation and quality (Dutt et al., 2004; Hua et al., 2009; Tan et al., 2013; Feng et al., 2015; Hande et al., 2017).

Several investigations show that the transcription level of structural flavonoid biosynthesis genes is higher in brown than in white cotton fibre. Transcription analysis reveals that the genes related to the PAs biosynthesis pathway from phenylalanine to flavan-3-ols were upregulated in brown cotton (Xiao et al., 2014). The expression level of main structural genes such as *GhCHI*,

GhF3H, *GhDFR*, *GhANS* and *GhANR* at 8, 16, 20 and 30 DPA was higher in brown compared with white and green cotton fibre (Xiao et al., 2007; Malik et al., 2015). Similar differences in the transcriptional activity of *GhPAL, GhCHI, GhDFR, Gh3GT, GhANR* (Li et al., 2013) and *GhC4H, GhCHS, GhF3'H, GhF3'5'H* genes (Feng et al., 2013) in white and brown cotton were noticed.

The high transcription level of *GhANS* and *GhANR* leads to anthocyanidin formation from leucoanthocyanidin (Peng et al., 2012; Malik et al., 2015). Since the expression of the identified genes play a key role in the PA biosynthesis pathway, it might be possible to control PA accumulation and, as a result, to control shades of brown colour during fibre development (Malik et al., 2015).

The transcription factors in PAs biosynthesis. Flavonoid biosynthesis is controlled by many transcription factors (TFs) such as R2R3-MYB-type factors, basic helix–loop–helix (bHLH) and WD40 repeat (Koes et al., 2005; Qi et al., 2011; Lloyd et al., 2017). These TFs could form a MBW complex that has a determinant role in the regulation of pigment biosynthesis during the flavonoid pathway (Baudry, Caboche, and Lepiniec, 2006). Using *Arabidopsis* as an experimental model, it was shown that R2R3-MYB proteins activate early anthocyanins and PA biosynthetic genes (EBGs) (Fig. 2). The expression of late biosynthetic genes (LBGs) is controlled by MYB-bHLH-WD40 repeat (MBW) complex (Fig. 2) (Li, 2014).

It was reported that in *Arabidopsis* the R2R3-MYB-coding gene *AtTT2* (TRANSPARENT TESTA 2) initiates the PAs biosynthesis (Nesi et al., 2001). Two homologous candidate genes — *GhMYB10* and *GhMYB36* (later *GhTT2-3A*) — were isolated from tetraploid cotton using BLAST analysis and related to R2R3-type MYB transcription factors (Table 2) (Lu, Roldan, and Dixon, 2017). Transactivation assays shown that these genes are the part of the regulative MBW complex that could activate *LAR* and *ANR* promoters (Albert et al., 2014; Jun et al., 2015; Li, Chen et al., 2016; Li, Dong et al., 2016; Liu, Jun, and Dixon, 2014). The simultaneous presence of both proved that *MYB*s from the *G. hirsutum* genome lead to more effective PAs biosynthesis (Lu et al., 2017).

The *GhTT2-A07* gene (later *GhTT2-3A*), which is related to brown colour cotton fibre as *Lc1* locus, was identified (Table 2) (Hinchliffe et al., 2016). Using transgenic analysis and expression levels, the *GhTT2-3A* gene has been shown to regulate the formation of brown fibres (Yan et al., 2018).

The *Arabidopsis* bHLH-type gene *AtTT8* (TRANSPARENT TESTA 8) is transcriptionally activated by *AtTT2*; it has a further synergic cooperative influence on the PAs structural genes and enhances their expression level (Baudry et al., 2004; Koes et al., 2005; Baudry, Caboche, and Lepiniec, 2006; Shangguan, Yang, Zhang, and Wang, 2016). Similar observations were reported for

Table 2. Regulatory genes which are involved in PAs biosynthesis pathway in cotton

G. hirsutum gene	Chromosome; Gene ID (CottonFGD*)	The protein family	References
GhMYB10	A06; Gohir.A06G075700 D06; Gohir.D06G074700	MYB domain	Lu, Roldan, and Dixon, 2017
GhTT2-3A (GhMYB36)	A07; Gh_A07G2341 D07; Gh_D07G0169		Hinchliffe et al., 2016
GhbHLH130D	A11; GH_A11G1273 D11; GH_D11G1302	bHLH domain	Yan et al., 2018
GhTTG1	A05; Gohir.A05G415900	WD-repeat	Humphries et al., 2005
GhTTG3	D04; Gohir.D04G000300		

* CottonFGD — Cotton Functional Genomics Database

GhTT2-3A and *GhbHLH130D* (*AtTT8* ortholog) cotton genes — it was established that *GhTT2-3A* activates the *GhbHLH130D* gene (Table 2) (Gong et al., 2018). Dual-luciferase assays demonstrated that the *GhTT2-3A* and *GhbHLH130D* genes affect the PAs biosynthesis structural genes, *GhANR* and *GhLAR*, and consequently induce PAs accumulation in cotton fibre (Yan et al., 2018).

It was established that WD40-coding gene *AtTTG1* (TRANSPARENT TESTA GLABRA 1) in *Arabidopsis* is responsible for seed coat production, root hair development and control of anthocyanins biosynthesis (Oppenheimer et al., 1991; Galway et al., 1994; Liu, Osbourn, and Ma, 2015). Among four *GhTTG1-GhTTG4* genes known to be involved in the cotton fibre growing initiation, *GhTTG1* and *GhTTG3* demonstrate close sequence similarity to the known *AtTTG1* anthocyanin regulatory gene (Table 2) (Walker et al., 1999; Mehboob-ur-Rahman et al., 2012; B. Liu, Zhu, and Zhang, 2015). Furthermore, a transient analysis determined that the *GhTTG1* and *GhTTG3* genes were able to complement the purple pigment in anthocyanin-deficit *ttg1 Arabidopsis* mutants, in contradistinction to *GhTTG2* and *GhTTG4* (Humphries et al., 2005; Marinova et al., 2007). Thus, WD-repeat *GhTTG1* and *GhTTG3* genes play a significant role during cotton fibre formation and in PAs biosynthesis regulation (Humphries et al., 2005).

Other genes. Structural genes involved in PAs modification, transport and oxidation processes have also been identified and investigated in the *G. hirsutum* genome. In particular, the *TT12* gene encodes proteins related to MATE (MULTIDRUG AND TOXIN EXTRUSION) family proteins which are responsible for the transfer of epicatechin-3'-O-glucoside, a precursor of PAs, into a central vacuole (Marinova and Pourcel, 2007; Zhao and Dixon, 2009). The *GhTT12* gene was studied using genetic engineering manipulations (Gao et al., 2016). A full-length cDNA of *TT12* from brown

cotton fibre was cloned in *N. tabacum* to study the expression level. The results have confirmed the potential participation of the *GhTT12* gene in PAs transportation (Gao et al., 2016).

Phylogenetic analysis supposes that cotton's MATE proteins are merged in the group with the TT12-like MATE transporters in *Arabidopsis* participating in PA transportation from the cytoplasmic matrix into the vacuole (Xu et al., 2018). The *GhMATE12*, *GhMATE16* and *GhMATE38* genes identified using phylogenetic analysis share a similar transcription activity with the *GhTT12* gene at specific development stages. As a result, detected genes could participate in transportation of PAs into the vacuole along with *GhTT12* (Xu et al., 2018). Besides, it was shown that the transcript level of the *GhMATE1a* and *GhMATE1b* genes was higher in brown than in white cotton (Feng et al., 2014). The highest value was enriched to 21 DPA for the *GhMATE1a* gene and to 27 DPA in the case of *GhMATE1b* (Feng et al., 2014).

It is known that Aha10 (H^+-APTase isoform 10) belongs to p-type ATPase family plasma membrane H^+ pump and is involved in the PAs accumulation process in the vacuole (Baxter et al., 2005; Appelhagen et al., 2015). *Aha10* mutants are characterized by the formation of a large number of small vacuoles instead of one big central vacuole. In addition, production of final metabolites in the flavonoid biosynthesis pathway such as PAs, anthocyanidins and epicatechins breaks down while the primary products are present (Baxter et al., 2005).

The *AtTT10* gene of *Arabidopsis* is responsible for a laccase-like 15 enzyme (LAC15) synthesis, involved in conversion of colourless epicatechin monomer and oligomers to brown oxidized tannins in the seed coat (Pourcel et al., 2005; Lloyd et al., 2017). The gene orthologue to *AtTT10* was identified in *G. hirsutum* (Hinchliffe et al., 2016).

The *AtTT13* gene from *Arabidopsis* is responsible for PA production in the seed coat endothelium (Gonzalez et al., 2016; Lloyd et al., 2017). The *GhTT13* gene, an orthologue to *AtTT13*, was detected in upland cotton *G. hirsutum* (Hinchliffe et al., 2016).

Thus, the control of PAs biosynthesis is controlled by factors which are also associated with anthocyanin biosynthesis. The main structural genes, as well as the components of the MBW complex controlling the synthesis of PAs, are currently discovered.

3. Caffeic acid derivates in green cotton fibre

Caffeic acid (CA) represents a cross-linked phenolic polymer that derived from three types of phenylpropane units: coniferyl, sinapyl and *p*-coumaryl alcohol (Fig. 3) (Schmutz et al., 1994; Ma et al., 2016a). CA is involved in lignin biosynthesis, which is a main structural component of the secondary cell wall in higher plants (Riley and Kolattukudy, 1975; Liu, Luo, and Zheng, 2018). It was reported that the secondary cell wall of cotton fibre contains mainly cellulose (about 94 %) and 0.4–1 % lignin (Fan et al., 2009).

It was shown that CA derivatives are responsible for the formation of green coloured cotton fibre: about 70 % of ω-hydroxydocosanoic acid and 25 % of docosanedoic acid were isolated from green cotton fibre, compared with 0.5 % in white fibre (Schmutz et al., 1993, 1994). Using an ultra-violet spectroscopy and a nuclear-magnetic-resonance (^{1}H-NMR) spectroscopy it has been shown that glycerol, CA and esterified CA constitute the main part of the wax fraction in green cotton fibre (Schmutz et al., 1993).

Green cotton fibre demonstrates the significant increase in UV-absorption properties — the radical-scavenging activity grate was approximately 20 times more than white cotton fibre due to the accumulation of CA derivatives (Re et al., 1999; Ma et al., 2016a; Masek, 2016). A decrease in the ability of radical scavenging capacities is observed under alkaline solution treatment of green and brown fibre cotton due to enchasing of fibre hydrophilic properties leading to barrier-free penetration of free radicals to the active component inside the fibre. The reason is alkaline hydrolysis of the ester bond and the release of water-soluble CA (Ma et al., 2016b). Moreover, the significant changes in fibre colour have been observed after alkaline treatment (Hinchliffe et al., 2015). This emphasizes the essential role of CA and its derivatives in fibre pigmentation and indicates the necessity of the phenylpropanoid pathway in colouration of green cotton fibre (Fan et al., 2009).

SUBERIZATION OF GREEN COTTON

Suberin and cutin biopolymers usually occur on the surface of higher plants — suberin accumulates into inner cell-wall and cutin is found at the surface of epidermal cells (Moire et al., 1999; Nawrath, 2002; Graça, 2015). These compounds protect plant leaves from adverse environmental factors (Kolattukudy, 1980; Yatsu et al., 1983). Suberin is formed from poly-functional fatty acids that are bound by ester bonds with glycerol (Kolattukudy, 1980; Matzke and Riederer, 1991; Graça, 2015) Epidermal cells of cotton seeds with green fibres produce both cutin and suberin (Ryser et al., 1983; Yatsu et al., 1983; Schmutz et al., 1993; Stankovič Elesini et al., 2002).

Suberization of the cotton seed coat is characteristic feature of *Gossypium* species (Ryser et al., 1985). However, a suberin layer has not been found in the seed coat of brown and white cotton fibres (Ryser et al., 1983, 1985; Schmutz et al., 1993). Multiple concentric rings of laminar ultrastructure similar to the suberin layer in the cell wall of green cotton fibre were demonstrated using electron microscopy

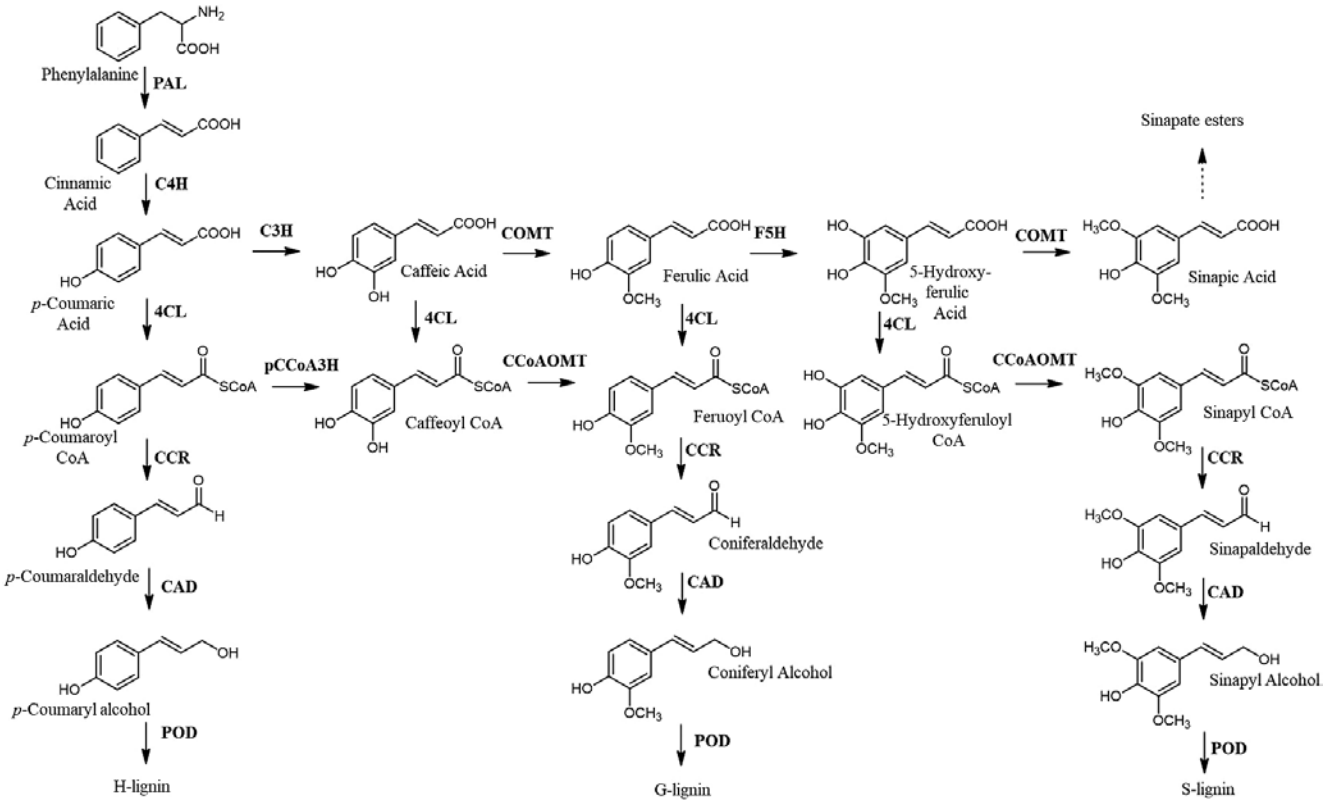

Fig. 3. Phenylpropanoid pathway of caffeic, ferulic and sinapic acids biosynthesis (modified from Humphreys, et al. 1999). The involved enzymes: PAL, phenylalanine ammonia-lyase; C4H, cinnamate 4-hydroxylase; C3H, *p*-Coumarate 3-hydroxylase; 4CL, 4-(hydroxy)cinnamoyl CoA ligase; COMT, caffeic acid 5- hydroxyferulic acid O-methyltransferase; F5H, ferulate 5-hydroxylase; pCCoA3H, *p*-Coumaroyl CoA 3-hydroxylase; CCoAOMT, caffeoyl CoA O-methyltransferase; CCR, cinnamoyl CoA reductase; CAD, cinnamoyl alcohol dehydrogenase; POD, peroxidase.

(Yatsu et al., 1983). Simultaneously, the white fibre cotton cell wall had a thin outer cuticle (Yatsu et al., 1983).

Phenylalanine ammonium-lyasa is recognized to be an important agent in converting phenylalanine into trans-cinnamic acid. Results of an experiment with the addition of 2-aminoindan-2-phosphonic acid, known as an inhibitor of phenylalanine ammonium-lyasa, to cultivating ovules have shown that an esterified CA is covalently linked with suberin and plays a significant role in the attachment of suberin to a cellulosic secondary wall (Schmutz et al., 1993). Moreover, green fibre colour could not be observed if the suberin formation is inhibited during tissue development.

There are some studies describing an accurate chemical structure of aliphatic components of the suberin layer (Schmutz et al., 1993; Ma et al., 2016a; Feng et al., 2017). It was demonstrated by reverse-phase analytical high-pressure liquid chromatography (HPLC) that extracted yellow pigment from the wax of green fibre cotton contains two major components and several minor ones. Namely, 22-O-caffeoyl-22-hydroxydocosanoic acid, glycerol ester, and 22-O-caffeoyl-22-hydroxydocosanoic acid were identified as main components of yellow pigment extract. According to the ultra-violet visible spectroscopy data, the minor components could be referred to CA derivatives (Ma et al., 2016a). Also, several fatty acids including dimethyl-2-hydroxysuccinate monoglyceride and heptacosanoic acid monoglyceride were isolated from green cotton fibre (Schmutz et al., 1993; Ma et al., 2016a; Feng et al., 2017). The structures were established by NMR spectroscopy on nuclei ^{1}H and ^{13}C, H-H Correlation Spectroscopy (H-COSY), Matrix Assisted Laser Desorption Ionization Time-of-Flight mass spectrometry (MALDI-TOF MS) and were coincided with earlier data (Schmutz et al., 1994, Torresdepinedo et al., 2007; Feng et al.,, 2017). It was found that from the isolated fatty acids, 22-O-caffeoyl-22-hydroxydocosanoic acid and 22-O-caffeoyl-22-hydroxymonodocosanoin (yellowish and a green-yellow powder, respectively) were demonstrated to be responsible for pigmentation (Feng et al., 2017). It was observed that the higher the concentration of 22-O-caffeoyl-22-hydroxymonodocosanoin, the deeper green colour of the cotton fibre (Feng et al., 2017).

STRUCTURAL GENES RELATED TO CA AND ITS DERIVATIVES BIOSYNTHESIS

CA biosynthesis is based on the phenylpropanoid pathway (Fig. 3) (Humphreys, Hemm, and Chapple, 1999; Boerjan, Ralph, and Baucher, 2003; Heldt and Heldt,

2005; Fan et al., 2009; Li et al., 2015). Identified genes involved in the CA biosynthesis pathway for *G. hirsutum* are listed in Table 1.

Trans-hydroxycinnamic acid is a product of the first stage catalysed by PAL, which transformed into *p*-coumaric acid due to para hydroxylation with C4H (Fig. 3). It should be noted that *p*-coumaric acid is a main precursor of other natural phenols (Boerjan et al., 2003; Ramawat and Mérillon, 2013; Shahidi and Yeo, 2018). Addition of OH-group is catalysed by *p*-coumarate 3-hydroxylase (C3H) at the ortho-position with CA formation, which serves as an initial compound in ferulic and synapic acids biosynthesis (Fig. 3) (Heldt and Heldt, 2005).

Gyaiacyl lignin (G-lignin) consists of such monolignols as coniferyl alcohol, syringine lignin (S-lignin) and *p*-Hydroxyphenyl lignin (H-lignin), formed from sinapyl alcohol and *p*-Coumaryl alcohol through the phenylpropanoid pathway, respectively (Fig. 3) (Humphreys et al., 1999; Boerjan et al., 2003; Vanholme et al., 2010; Li, Pu, and Ragauskas, 2016). The ratio of monolignols affects the physical properties of lignin polymer (Macmillan et al., 2013).

From two phenylalanine ammonium-lyase genes, *GhPAL* and *GhPAL2*, the expression of *GhPAL* has a higher level during initiation of fibre secondary cell wall thickening (Ling Fan, Wen-Ran Hu, 2012; Feng et al., 2014; Qin et al., 2017). It has been shown that the *GhC4H1* and *GhC4H2* genes involved in lignin and flavonoid biosynthesis processes are expressed in developing cotton fibres (Ling Fan and Wen-Ran Hu, 2012).

Gh4CL1-GhCL4 cotton genes encode 4-Coumarate CoA ligases (4CL), which are responsible for converting *p*-coumaric acid to *p*-Coumaroyl CoA and its derivatives such as ferulic, caffeic and sinapic acids into CoA esters (Figs. 2, 3). These compounds take part in the phenylpropanoid pathway, providing the plant secondary metabolites formation (Hamberger and Hahlbrock, 2004; Ling Fan, Wen-Ran Hu, 2012; Li et al., 2015). Four isoforms of the *Gh4CL1-GhCL4* genes were identified in the genome of white and green *G. hirsutum* L. The expression profiles studding of *4CL* structural genes in white and green cotton fibre were assayed by RT-PCR (Feng et al., 2017). *Gh4CL3* and *Gh4CL4* demonstrate a higher transcript level in roots and hypocotyls than in flowers and leaves and have a similar expression profile in white and green fibre gradually increased to 24 days of growing. *Gh4CL1* expression decreased with fibre development in white and green cotton. The highest expression level among all considered genes was observed in *Gh4CL2* in green fibre (Feng et al., 2017).

The catalytic properties of proteins related to *4CL* cotton structural genes displayed that Gh4CL2 has the highest turnover rate for caffeic, cinnamate and ferulate in comparison to Gh4CL3 and Gh4CL4 (Feng et al., 2017). Gh4CL1 showed a low turnover rate also, but demonstrated a preference for 4-coumarat synthesis (Feng et al., 2017). Considering the results of expression level studding and enzyme activity of corresponding genes, it can be proposed that the *Gh4CL2* gene is involved in the metabolism of CA derivatives and is responsible for pigmentation of green cotton fibres (Ling Fan and Wen-Ran Hu, 2012).

It was found that among seven *GhCAD1-GhCAD7* genes similar to *AtCAD5* in *A. thaliana*, only the *GhCAD6* gene showed up-regulation during cotton fibre development. This result displays the involvement of this gene in the phenylpropanoid pathway (Fan et al., 2009; Ling Fan and Wen-Ran Hu, 2012).

Thus, the nucleotide sequences of most genes involved in CA derivatives biosynthesis and modification have not yet been identified.

Conclusion

The article provides comprehensive knowledge of historical aspects and describes the various approaches associated with the research of phenolic pigments isolated from coloured cotton. Structural and regulatory genes involved in PAs and CA biosynthesis have been characterized in the current review. Biosynthesis of these plant phenolic compounds occurs via the phenylpropanoid pathway. The metabolism of PAs is most fully studied at the genetic level. However, there are no studies concerning transcriptional regulation of CA biosynthesis. In addition, we have a limited understanding of the pleiotropic influence of the genes responsible for colour on the fibre quality attributes. Further study of the genes related to PAs and CA biosynthesis in combination with available data will make it possible to reconstruct the complete genetic regulatory network of the biosynthesis of these compounds.

Acknowledgments

We thank Chukhina I.G., PhD, Leading Researcher (VIR) and Podolnaya L.P., PhD, Leading Researcher (VIR) for providing samples from Herbarium collection. This work was conducted within the framework of the state task regarding Theme № 0481-2019-0001 "Genomic and postgenomic technologies for searching for new genetic markers of economically important traits and new allelic variants of agricultural valuable genes in the gene pool of cultivated plants and their wild relatives".

References

Abdurakhmonov, I. Y., Ayubov, M. S., Ubaydullaeva, K. A., Buriev, Z. T., Shermatov, S. E., Ruziboev, H. S., Shapulatov, U. M., Saha, S., Ulloa, M., Yu, J. Z., Percy, R. G., Devor, E. J., Sharma, G. C., Sripathi, V. R., Kumpatla, S. P., van der Krol, A., Kater, H. D., Khamidov, K., Salikhov, S. I., Jenkins, J. N., Abdukarimov, A., and Pepper, A. E. 2016. RNA interference for functional genomics and improvement of cotton (*Gossypium* sp.). *Frontiers in Plant Science* 7(202):1–17. https://doi.org/10.3389/fpls.2016.00202

Albert, N. W., Davies, K. M., Lewis, D. H., Zhang, H., Montefiori, M., Brendolise, C., Boase, M. R., Ngo, N., Jameson, P. E., and Schwinn, K. E. 2014. A conserved network of transcriptional activators and repressors regulates anthocyanin pigmentation in eudicots. *The Plant Cell* 26(3):962–980. https://doi.org/10.1105/tpc.113.122069

Alonso-Amelot, M. E., Oliveros, A., and Calcagno-Pisarelli, M. P. 2004. Phenolics and condensed tannins in relation to altitude in neotropical *Pteridium* spp. *Biochemical Systematics and Ecology* 32(11):969–981. https://doi.org/10.1016/j.bse.2004.03.005

Appelhagen, I., Nordholt, N., Seidel, T., Spelt, K., Koes, R., Quattrochio, F., Francesca, S., and Weisshaar, B. 2015. TRANSPARENT TESTA 13 is a tonoplast P 3A -ATPase required for vacuolar deposition of proanthocyanidins in *Arabidopsis thaliana* seeds. *The Plant Journal* 82(5):840–849. https://doi.org/10.1111/tpj.12854

Balls, W. L. 1908. Mendelian studies of Egyptian cotton. *The Journal of Agricultural Science* 2(04):346. https://doi.org/10.1017/S0021859600000654

Baudry, A., Caboche, M., and Lepiniec, L. 2006. TT8 controls its own expression in a feedback regulation involving TTG1 and homologous MYB and BHLH factors, allowing a strong and cell-specific accumulation of flavonoids in *Arabidopsis thaliana*. *The Plant Journal* 46(5):768–779. https://doi.org/10.1111/j.1365-313X. 2006.02733.x

Baudry, A., Heim, M. A., Dubreucq, B., Caboche, M., Weisshaar, B., and Lepiniec, L. 2004. TT2, TT8, and TTG1 synergistically specify the expression of BANYULS and proanthocyanidin biosynthesis in *Arabidopsis thaliana*. *The Plant Journal* 39(3):366–380. https://doi.org/10.1111/j.1365-313X. 2004.02138.x

Baxter, I. R., Young, J. C., Armstrong, G., Foster, N., Bogenschutz, N., Cordova, T., Peer, W. A., Hazen, S. P., Murphy, A. S., and Harper, J. F. 2005. A plasma membrane H+-ATPase is required for the formation of proanthocyanidins in the seed coat endothelium of *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences* 102(7):2649–2654. https://doi.org/10.1073/pnas.0406377102

Boerjan, W., Ralph, J., and Baucher, M. 2003. Lignin biosynthesis. *Annual Review of Plant Biology* 54(1):519–546. https://doi.org/10.1146/annurev.arplant.54.031902.134938

Bogs, J., Downey, M. O., Harvey, J. S., Ashton, A. R., Tanner, G. J., and Robinson, S. P. 2005. Proanthocyanidin synthesis and expression of genes encoding leucoanthocyanidin reductase and anthocyanidin reductase in developing grape berries and grapevine leaves. *Plant Physiology* 139(2):652–663. https://doi.org/10.1104/pp.105.064238

Debeaujon, I., Nesi, N., Perez, P., Devic, M., Grandjean, O., Caboche, M., and Lepiniec, L. 2003. Proanthocyanidin-accumulating cells in *Arabidopsis* testa: regulation of differentiation and role in seed development. *The Plant Cell Online* 15(11):2514–2531. https://doi.org/10.1105/tpc.014043

Dinesh, K. A., Phundan, S., Mukta, Ch., Shaikh, A. J. 2003. Cottonseed oil quality, utilization and processing. *Cicr Technical Bulletin* 25:1–16.

Dutt, Y., Wang, X. D., Zhu, Y. G., and Li, Y. Y. 2004. Breeding for high yield and fibre quality in coloured cotton. *Plant Breeding* 123(2):145–151. https://doi.org/10.1046/j.1439-0523.2003.00938.x

Endrizzi, J. E. and Taylor, T. 1968. Cytogenetic studies of N Lc1 Yg2 R2 marker genes and chromosome deficiences in cotton. *Genetical Research* 12(03):295–304. https://doi.org/10.1017/S0016672300011885

Ershik, O. A. and Buzuk, G. N. 2009. Identification of the degradation products of proanthocyanidines in root rhizomes of the marsh cinquefoil. *Pharmaceutical Chemistry Journal* 43(7):406–407. https://doi.org/10.1007/s11094-009-0315-y

Fan, L., Shi, W.-J., Hu, W.-R., Hao, X.-Y., Wang, D.-M., Yuan, H., and Yan, H.-Y. 2009. Molecular and biochemical evidence for phenylpropanoid synthesis and presence of wall-linked phenolics in cotton fibers. *Journal of Integrative Plant Biology* 51(7):626–637. https://doi.org/10.1111/j.1744-7909.2009.00840.x

Feng, H., Guo, L., Wang, G., Sun, J., Pan, P., He, S., Zhu, H., Sun, J., and Du, X. 2015. The negative correlation between fiber color and quality traits revealed by QTL analysis. *PLoS ONE* 10(6):e0129490. https://doi.org/10.1371/journal.pone.0129490

Feng, H., Li, Y., Wang, S., Zhang, L., Liu, Y., Xue, F., Sun, Y., Wang, Y., and Sun, J. 2014. Molecular analysis of proanthocyanidins related to pigmentation in brown cotton fibre (*Gossypium hirsutum* L.). *Journal of Experimental Botany* 65(20):5759–5769. https://doi.org/10.1093/jxb/eru286

Feng, H., Tian, X., Liu, Y., Li, Y., Zhang, X., Jones, B. J., Sun, Y., and Sun, J. 2013. Analysis of flavonoids and the flavonoid structural genes in brown fiber of upland cotton. *PLoS ONE* 8(3):e58820. https://doi.org/10.1371/journal.pone.0058820

Feng, H., Yang, Y., Sun, S., Li, Y., Zhang, L., Tian, J., Zhu, Q., Feng, Z., Zhu, H., and Sun, J. 2017. Molecular analysis of caffeoyl residues related to pigmentation in green cotton fibers. *Journal of Experimental Botany* 68(16):4559–4569. https://doi.org/10.1093/jxb/erx281

French A. D. and Kim H. J. 2018. Cotton fiber structure. Cotton fiber: physics, chemistry and biology, edited by D. D. Fang. Springer, Berlin, pp. 13–39. https://doi.org/10.1007/978-3-030-00871-0_2

Galway, M. E., Masucci, J. D., Lloyd, A. M., Walbot, V., Davis, R. W., and Schiefelbein, J. W. 1994. The TTG gene is required to specify epidermal cell fate and cell patterning in the *Arabidopsis* root. *Developmental Biology* 166:740–754. https://doi.org/10.1006/dbio.1994.1352

Gao, J.-S., Wu, N., Shen, Z.-L., Lv, K., Qian, S.-H., Guo, N., Sun, X., Cai, Y.-P., and Lin, Y. 2016. Molecular cloning, expression analysis and subcellular localization of a Transparent Testa 12 ortholog in brown cotton (*Gossypium hirsutum* L.). *Gene* 576(2):763–769. https://doi.org/10.1016/j.gene.2015.11.002

Gong, W., Du, X., Jia, Y., and Pan, Z. 2018. Color cotton and its utilization in China. Cotton fiber: physics, chemistry and biology, edited by D. D. Fang. Springer, Berlin, pp. 117–132. https://doi.org/10.1007/978-3-030-00871-0_6

Gong, W., He, S., Tian, J., Sun, J., Pan, Z., Jia, Y., Sun, G., and Du, X. 2014. Comparison of the transcriptome between two cotton lines of different fiber color and quality. *PLoS ONE* 9(11):e112966. https://doi.org/10.1371/journal.pone.0112966

Gonzalez, A., Brown, M., Hatlestad, G., Akhavan, N., Smith, T., Hembd, A., Moore, J., Montes, J., Mosley, T., Resendez, J., Nguyen, H., Wilson, L., Campbell, A., Sudarshan, D., and Lloyd, A. 2016. TTG2 controls the developmental regulation of seed coat tannins in *Arabidopsis* by regulating vacuolar transport steps in the proanthocyanidin pathway. *Developmental Biology* 419(1):54–63. https://doi.org/10.1016/j.ydbio.2016.03.031

Graça, J. 2015. Suberin: the biopolyester at the frontier of plants. *Frontiers in Chemistry* 3(62):1–11. https://doi.org/10.3389/fchem.2015.00062

Gu, H. 2005. Research on the improvement of the moisture absorbency of naturally self-coloured cotton. *Journal of the Textile Institute* 96(4):247–250. https://doi.org/10.1533/joti.2005.0007

Hamberger, B. and Hahlbrock, K. 2004. The 4-coumarate:CoA ligase gene family in *Arabidopsis thaliana* comprises one rare, sinapate-activating and three commonly occurring isoenzymes. *Proceedings of the National Academy of Sciences* 101(7):2209–2214. https://doi.org/10.1073/pnas.0307307101

Hande, A. S., Katageri, I. S., Jadhav, M. P., Adiger, S., Gamanagatti, S., Padmalatha, K. V., Dhandapani, G., Kanakachari, M., Kumar, P. A., and Reddy, V. S. 2017. Transcript profiling of genes expressed during fibre development in diploid cotton (*Gossypium arboreum* L.). *BMC Genomics* 18(675):1–15. https://doi.org/10.1186/s12864-017-4066-y

Harland, S. C. 1932. The genetics of cotton. *Journal of Genetics* 25(3):261–270. https://doi.org/10.1007/BF02984590

He, F., Pan, Q.-H., Shi, Y., and Duan, C.-Q. 2008. Biosynthesis and genetic regulation of proanthocyanidins in plants. *Molecules* 13(10):2674–2703. https://doi.org/10.3390/molecules13102674

Heldt, H.-W. and Heldt, F. 2005. Phenylpropanoids comprise a multitude of plant secondary metabolites and cell wall components. Plant Biochemistry (Third Edition), Academic Press, pp. 435–454. https://doi.org/10.1016/B978-012088391-2/50019-5

Hinchliffe, D., Condon, B., Delhom, C. D., Chang, S., Montalvo, J., Madison, C., Reynolds, M., VonHoven, T., and Cintrón, M. S. 2015. Physical and combustion properties of nonwoven fabrics produced from conventional and naturally colored cottons. *Textile Research Journal* 85(16):1666–1680. https://doi.org/10.1177/0040517515573410

Hinchliffe, D. J., Condon, B. D., Thyssen, G., Naoumkina, M., Madison, C. A., Reynolds, M., Delhom, C. D., Fang, D. D., Li, P., and McCarty, J. 2016. The GhTT2_A07 gene is linked to the brown colour and natural flame retardancy phenotypes of Lc1 cotton (*Gossypium hirsutum* L.) fibres. *Journal of Experimental Botany* 67(18):5461–5471. https://doi.org/10.1093/jxb/erw312

Hua, S. J., Yuan, S. N., Shamsi, I. H., Zhao, X. Q., Zhang, X. Q., Liu, Y. X., et al. 2009. A comparison of three isolines of cotton differing in fiber color for yield, quality, and photosynthesis. *Crop Science* 49:983–989. https://doi.org/10.2135/cropsci2008.06.0371

Humphreys, J. M., Hemm, M. R., and Chapple, C. 1999. New routes for lignin biosynthesis defined by biochemical characterization of recombinant ferulate 5-hydroxylase, a multifunctional cytochrome P450-dependent monooxygenase. *Proceedings of the National Academy of Sciences* 96(18):10045–10050. https://doi.org/10.1073/pnas.96.18.10045

Humphries, J. A., Walker, A. R., Timmis, J. N., and Orford, S. J. 2005. Two WD-Repeat genes from cotton are functional homologues of the *Arabidopsis thaliana* TRANSPARENT TESTA GLABRA1 (TTG1) gene. *Plant Molecular Biology* 57(1):67–81. https://doi.org/10.1007/s11103-004-6768-1

Hutchinson, J. B. and Silow, R. A. 1939. Gene symbols for use in cotton genetics. *Journal of Heredity* 30(10):461–464. https://doi.org/10.1093/jhered/30.10.461

Jun, J. H., Liu, C., Xiao, X., and Dixon, R. A. 2015. The transcriptional repressor MYB2 regulates both spatial and temporal patterns of proanthocyandin and anthocyanin pigmentation in *Medicago truncatula*. *The Plant Cell* 27(10):2860–2879. https://doi.org/10.1105/tpc.15.00476

Koes, R., Verweij, W., and Quattrocchio, F. 2005. Flavonoids: a colorful model for the regulation and evolution of biochemical pathways. *Trends in Plant Science* 10(5):236–42. https://doi.org10.1016/j.tplants.2005.03.002

Kolattukudy, P. E. 1980. Biopolyester membranes of plants: cutin and suberin. *Science* 208(4447):990–1000. https://doi.org/10.1126/science.208.4447.990

Kottur, G. L. 1923. Studies on inheritance in cotton I. History of a cross between *Gossypium herbaceum* and *G. neglectum*. *Memories Dept. Agri. Indian Bot.* 12:71–133.

Li, M., Pu, Y., and Ragauskas, A. J. 2016. Current understanding of the correlation of lignin structure with biomass recalcitrance. *Frontiers in Chemistry* 4(45):1–8. https://doi.org/10.3389/fchem.2016.00045

Li, P., Chen, B., Zhang, G., Chen, L., Dong, Q., Wen, J., Mysore, K. S., and Zhao, J. 2016. Regulation of anthocyanin and proanthocyanidin biosynthesis by *Medicago truncatula* BHLH transcription factor MtTT8. *New Phytologist* 210(3):905–921. https://doi.org/10.1111/nph.13816

Li, P., Dong, Q., Ge, S., He, X., Verdier, J., Li, D., and Zhao, J. 2016. Metabolic engineering of proanthocyanidin production by repressing the isoflavone pathways and redirecting anthocyanidin precursor flux in legume. *Plant Biotechnology Journal* 14(7):1604–1618. https://doi.org/10.1111/pbi.12524

Li, S. 2014. Transcriptional control of flavonoid biosynthesis. *Plant Signaling & Behavior* 9(1):e27522. https://doi.org/10.4161/psb.27522

Li, Y.-J., Zhang, X.-Y., Wang, F.-X., Yang, C.-L., Liu, F., Xia, G. X., and Sun, J. 2013. A comparative proteomic analysis provides insights into pigment biosynthesis in brown color fiber. *Journal of Proteomics* 78:374–88. https://doi.org/10.1016/j.jprot.2012.10.005

Li, Y., Kim, J. I., Pysh, L., and Chapple, C. 2015. Four isoforms of *Arabidopsis thaliana* 4-Coumarate: CoA ligase (4CL) have overlapping yet distinct roles in phenylpropanoid metabolism. *Plant Physiology* 169(4): 2409–2421. https://doi.org/10.1104/pp.15.00838

Ling, F., Hu, W.-R., Yang, Y., and Li, B. 2012. Plant special cell — cotton fiber 1. Botany, Dr. John Mworia (Ed.), pp. 211–226.

Liu, B., Zhu, Y., and Zhang, T. 2015. The R3-MYB gene GhCPC negatively regulates cotton fiber elongation. *PLoS ONE* 10(2):e0116272. https://doi.org/10.1371/journal.pone.0116272

Liu, C., Jun, J. H., and Dixon, R. A. 2014. MYB5 and MYB14 play pivotal roles in seed coat polymer biosynthesis in *Medicago truncatula*. *Plant Physiology* 165(4):1424–39. https://doi.org/10.1104/pp.114.241877

Liu, H., Luo, C., Song, W., Shen, H., Li, G., He, Z., Chen, W., Cao, Y., Huang, F., Tang, S., Hong, P., Zhao, E., Zhu, J., He, D., Wang, S., Huo, G., Liu, H. 2018. Flavonoid biosynthesis controls fiber color in naturally colored cotton. *PeerJ* 6:e4537. https://doi.org/10.7717/peerj.4537

Liu, J., Osbourn, A., and Ma, P. 2015. MYB transcription factors as regulators of phenylpropanoid metabolism in plants. *Molecular Plant* 8(5):689–708. https://doi.org/10.1016/j.molp.2015.03.012

Liu, Q., Luo, L., and Zheng, L. 2018. Lignins: biosynthesis and biological functions in plants. *International Journal of Molecular Sciences* 19(2):335. https://doi.org/10.3390/ijms19020335

Lloyd, A., Brockman, A., Aguirre, L., Campbell, A., Bean, A., Cantero, A., and Gonzalez, A. 2017. Advances in the MYB–bHLH–WD repeat (MBW) pigment regulatory model: addition of a WRKY factor and co-option of an anthocyanin MYB for betalain regulation. *Plant and Cell Physiology* 58(9):1431–1441. https://doi.org/10.1093/pcp/pcx075

Lu, N., Roldan, M., and Dixon, R. A. 2017. Characterization of two TT2-type MYB transcription factors regulating proanthocyanidin biosynthesis in tetraploid cotton, *Gossypium hirsutum*. *Planta* 246(2):323–335. https://doi.org/10.1007/s00425-017-2682-z

Ma, M., Hussain, M., Memon, H., and Zhou, W. 2016a. Structure of pigment compositions and radical scavenging activity of

naturally green-colored cotton fiber. *Cellulose* 23(1):955–963. https://doi.org/10.1007/s10570-015-0830-9

Ma, M., Luo, S., Hu, Z., Tang, Z., and Zhou, W. 2016b. Antioxidant properties of naturally brown-colored cotton fibers. *Textile Research Journal* 86(3):256–263. https://doi.org/10.1177/0040517515588270

Macmillan, C. P., Birke, H., Bedon, F., and Pettolino, F. A. 2013. Plant biochemistry and physiology lignin deposition in cotton cells — where is the lignin? 1(2):2–5. https://doi.org/10.4172/2329-9029.1000e106

Malik, W., Khan, A. A., Cheema, H. M. N., Aslam, U., Iqbal, M. Z., Qayyum, A., Yasmeen, A., and Bibi, N. 2015. Transcriptome analysis of pigment related genes in colored cotton. *International Journal of Agriculture and Biology* 17(1):205–210.

Marinova, K., Pourcel, L., Weder, B., Schwarz, M., Barron, D., Routaboul, J.-M., Debeaujon, I., and Klein, M. 2007. The *Arabidopsis* MATE transporter TT12 acts as a vacuolar flavonoid/H^+-antiporter active in proanthocyanidin-accumulating cells of the seed coat. *The Plant Cell* 19(6):2023–2038. https://doi.org/10.1105/tpc.106.046029

Masek, A., Chrzescijanska, E., and Latos, M. 2016. Determination of antioxidant activity of caffeic acid and p-coumaric acid by using electrochemical and spectrophotometric assays. *International Journal of Electrochemical Science* 11(12):10644–10658. https://doi.org/10.20964/2016.12.73

Matzke, K. and Riederer, M. 1991. A comparative study into the chemical constitution of cutins and suberins from *Picea abies* (L.) Karst., *Quercus robur* L., and *Fagus sylvatica* L. *Planta* 185(2):233–245. https://doi.org/10.1007/BF00194066

Mehboob-ur-Rahman, Shaheen, T., Tabbasam, N., Iqbal, M. A., Ashraf, M., Zafar, Y., and Paterson, A. H. 2012. Cotton genetic resources. A review. *Agronomy for Sustainable Development* 32(2):419–432. https://doi.org/10.1007/s13593-011-0051-z

Moire, L., Schmutz, A., Buchala, A., Yan, B., Stark, R. E., and Ryser, U. 1999. Glycerol is a suberin monomer. New experimental evidence for an old hypothesis. *Plant Physiology* 119(3):1137–1146. https://doi.org/10.1104/pp.119.3.1137

Nawrath, C. 2002. The biopolymers cutin and suberin. *The Arabidopsis Book* 1:e0021. https://doi.org/10.1199/tab.0021

Nechepurenko, I. V., Komarova, N. I., Gerasimova, Yu. V., Koval', V. V., Polovinka, M. P., Korchagina, D. V., and Salakhutdinov, N. F. 2009. Structure of oligomeric proanthocyanidines from *Hedysarum thienum* roots studied by thiolysis and MALDI-TOF MS. *Chemistry of Natural Compounds* 45(1):32–39. https://doi.org/10.1007/s10600-009-9216-2

Nesi, N., Jond, C., Debeaujon, I., Caboche, M., and Lepiniec, L. 2001. The *Arabidopsis* TT2 gene encodes an R2R3 MYB domain protein that acts as a key determinant for proanthocyanidin accumulation in developing seed. *The Plant Cell* 13(9):2099–2114. https://doi.org/10.1105/TPC.010098

Nimon, W. and Beghin, J. 1999. Are eco-labels valuable? Evidence from the apparel industry. *American Journal of Agricultural Economics* 81(4):801–811. https://doi.org/10.2307/1244325

Oppenheimer, D. G., Herman, P. L., Sivakumaran, S., Esch, J., and Marks, M. D. 1991. A myb gene required for leaf trichome differentiation in *Arabidopsis* is expressed in stipules. *Cell* 67: 483–493. https://doi.org/10.1016/0092-8674(91)90523-2

Panche, A. N., Diwan, A. D., and Chandra, S. R. 2016. Flavonoids: an overview. *Journal of Nutritional Science* 5(e47):1–15. https://doi.org/10.1017/jns.2016.41

Peng, Q.-Z., Zhu, Y., Liu, Z., Du, C., Li, K.-G., and Xie, D.-Y. 2012. An integrated approach to demonstrating the ANR pathway of proanthocyanidin biosynthesis in plants. *Planta* 236(3):901–918. https://doi.org/10.1007/s00425-012-1670-6

Pourcel, L., Routaboul, J.-M., Kerhoas, L., Caboche, M., Lepiniec, L., and Debeaujon, I. 2005. TRANSPARENT TESTA10 encodes a laccase-like enzyme involved in oxidative polymerization of flavonoids in *Arabidopsis* seed coat. *The Plant Cell* 17(11):2966–2980. https://doi.org/10.1105/tpc.105.035154

Prasad, D. 2000. Two A-type proanthocyanidins from *Prunus armeniaca* roots. *Fitoterapia* 71(3):245–253. https://doi.org/10.1016/S0367-326X(99)00165-3

Qi, T., Song, S., Ren, Q., Wu, D., Huang, H., Chen, Y., Fan, M., Peng, W., Ren, C., and Xie, D. 2011. The jasmonate-ZIM-domain proteins interact with the WD-Repeat/BHLH/MYB complexes to regulate jasmonate-mediated anthocyanin accumulation and trichome initiation in *Arabidopsis thaliana*. *The Plant Cell* 23(5):1795–1814. https://doi.org/10.1105/tpc.111.083261

Qin, Y., Wei, H., Sun, H., Hao, P., Wang, H., Su, J., and Yu, S. 2017. Proteomic analysis of differences in fiber development between wild and cultivated *Gossypium hirsutum* L. *Journal of Proteome Research* 16(8):2811–2824. https://doi.org/10.1021/acs.jproteome.7b00122

Ramawat, K. G. and Mérillon, J. M. 2013. Phenolic acids. Natural products. Edited by K. G. Ramawat and J.-M. Mérillon. Berlin, Heidelberg: Springer Berlin Heidelberg. 63:1952–1968. https://doi.org/10.1007/978-3-642-22144-6_64

Re, R., Pellegrini, N., Proteggente, A., Pannala, A., Yang, M., and Rice-Evans, C. 1999. Antioxidant activity applying an improved abts radical cation decolorization assay. *Free Radical Biology and Medicine* 26(9–10):1231–1237. https://doi.org/10.1016/S0891-5849(98)00315-3

Richmond, T. R. 1943. Inheritance of green and brown lint in upland cotton. *Agronomy Journal* 35(11):967–975. https://doi.org/10.2134/agronj1943.00021962003500110006x

Riley, R. G. and Kolattukudy, P. E. 1975. Evidence for covalently attached p-coumaric acid and ferulic acid in cutins and suberins. *Plant Physiology* 56(5):650–654. https://doi.org/10.1104/pp.56.5.650

Ryser, U. and Holloway, P. J. 1985. Ultrastructure and chemistry of soluble and polymeric lipids in cell walls from seed coats and fibres of *Gossypium* species. *Planta* 163(2):151–163. https://doi.org/10.1007/BF00393501

Ryser, U., Meier, H., and Holloway, P. J. 1983. Identification and localization of suberin in the cell walls of green cotton fibres (*Gossypium hirsutum* L., var. green lint). *Protoplasma* 117(3):196–205. https://doi.org/10.1007/BF01281823

Salunkhe, D. K., Jadhav, S. J., Kadam, S. S., and Chavan, J. K. 1983. Chemical, biochemical, and biological significance of polyphenols in cereals and legumes. *C R C Critical Reviews in Food Science and Nutrition* 17(3):277–305. https://doi.org/10.1080/10408398209527350

Schijlen, E. G. W. M., Ric de Vos, C. H., van Tunen, A. J., and Bovy, A. G. 2004. Modification of flavonoid biosynthesis in crop plants. *Phytochemistry* 65(19):2631–2648. https://doi.org/10.1016/j.phytochem.2004.07.028

Schmutz, A., Jenny, T., Amrhein, N., and Ryser, U. 1993. Caffeic acid and glycerol are constituents of the suberin layers in green cotton fibres. *Planta* 189(3):453–460. https://doi.org/10.1007/BF00194445

Schmutz, A., Jenny, T., and Ryser, U. 1994. A caffeoyl-fatty acid-glycerol ester from wax associated with green cotton fibre suberin. *Phytochemistry* 36(6):1343–1346. https://doi.org/10.1016/S0031-9422(00)89721-6

Shahidi, F. and Yeo, J. 2018. Bioactivities of phenolics by focusing on suppression of chronic diseases: a review.

International Journal of Molecular Sciences 19(6):1573. https://doi.org/10.3390/ijms19061573

Shangguan, X., Yang, C., Zhang, X., and Wang, L. 2016. Functional characterization of a basic helix-loop-helix (bHLH) transcription factor GhDEL65 from cotton (*Gossypium hirsutum*). *Physiologia Plantarum* 158(2):200–212. https://doi.org/10.1111/ppl.12450

Silow, R. A. 1944. The inheritance of lint colour in asiatic cottons. *Journal of Genetics* 46(1):78–115. https://doi.org/10.1007/BF02986695

Stankovič, E. U., Čuden A. P., and Richards, A. F. 2002. Study of the green cotton fibres. *Acta Chimica Slovenica* 49(4):815–833.

Takahashi, A., Ichihara, Y., Isagi, Y., and Shimada, T. 2010. Effects of acorn tannin content on infection by the fungus *Ciboria batschiana*. *Forest Pathology* 40(2):96–99. https://doi.org/10.1111/j.1439-0329.2009.00612.x

Tan, J., Tu, L., Deng, F., Hu, H., Nie, Y., and Zhang, X. 2013. A genetic and metabolic analysis revealed that cotton fiber cell development was retarded by flavonoid naringenin. *Plant Physiology* 162(1):86–95. https://doi.org/10.1104/pp.112.212142

Tan, J., Wang, M., Tu, L., Nie, Y., Lin, Y., and Zhang, X. 2013. The flavonoid pathway regulates the petal colors of cotton flower. *PLoS ONE* 8(8):e72364. https://doi.org/10.1371/journal.pone.0072364

Thomas, B. F. and ElSohly, M. A. 2016. Biosynthesis and pharmacology of phytocannabinoids and related chemical constituents. The Analytical Chemistry of Cannabis, Elsevier, pp. 27–41. https://doi.org/10.1016/B978-0-12-804646-3.00002-3

Torresdepinedo, A., Penalver, P., and Repezvictoria, I. 2007. Synthesis of new phenolic fatty acid esters and their evaluation as lipophilic antioxidants in an oil matrix. *Food Chemistry* 105(2):657–65. https://doi.org/10.1016/j.foodchem.2007.04.029

Vanholme, R., Demedts, B., Morreel, K., Ralph, J., and Boerjan, W. 2010. Lignin biosynthesis and structure. *Plant Physiology* 153(3):895–905. https://doi.org/10.1104/pp.110.155119

Walker, A. R., Davison, P. A., Bolognesi-Winfield, A. C., James, C. M., Srinivasan, N., Blundell, T. L., Esch, J. J., Marks, M. D., and Gray, J. C. 1999. The TRANSPARENT TESTA GLABRA1 locus, which regulates trichome differentiation and anthocyanin biosynthesis in *Arabidopsis*, encodes a WD40 repeat protein. *The Plant Cell* 11(7):1337–1349. https://doi.org/10.2307/3870753

Wang, L., Zhu, Y., Wang, P., Fan, Q., Wu, Y., Peng, Q.-Z., Xia, G.-X., and Wu, J.-H. 2016. Functional characterization of a dihydroflavanol 4-reductase from the fiber of upland cotton (*Gossypium hirsutum*). *Molecules* 21(2):32. https://doi.org/10.3390/molecules21020032

Wang, X., Howell, C. P., Chen, F., Yin, J., and Jiang, Y. 2009. Chapter 6 Gossypol-A polyphenolic compound from cotton plant. *Advances in Food and Nutrition Research* 58:215–263. https://doi.org/10.1016/S1043-4526(09)58006-0

Ware, J. O. 1932. Inheritance of lint colors in upland cotton. *Agronomy Journal* 24(7):550. https://doi.org/10.2134/agronj1932.00021962002400070008x

Xiao, Y.-H., Yan, Q., Ding, H., Luo, M., Hou, L., Zhang, M., Yao, D., Liu, H.-S., Li, X., Zhao, J., and Pei, Y. 2014. Transcriptome and biochemical analyses revealed a detailed proanthocyanidin biosynthesis pathway in brown cotton fiber. *PLoS ONE* 9(1):e86344. https://doi.org/10.1371/journal.pone.0086344

Xiao, Y.-H., Zhang, Z.-S., Yin, M.-H., Luo, M., Li, X.-B., Hou, L., and Pei, Y. 2007. Cotton flavonoid structural genes related to the pigmentation in brown fibers. *Biochemical and Biophysical Research Communications* 358(1):73–78. https://doi.org/10.1016/j.bbrc.2007.04.084

Xu, L., Shen, Z. L., Chen, W., Si, G. Y., Meng, Y., Guo, N., Sun, X., Cai, Y. P., Lin, Y., and Gao, J. S. 2018. Phylogenetic analysis of upland cotton MATE gene family reveals a conserved subfamily involved in transport of proanthocyanidins. *Molecular Biology Reports* 46:161–175. https://doi.org/10.1007/s11033-018-4457-4

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., and Xiao, Y. 2018. Up-regulation of GhTT2-3A in cotton fibres during secondary wall thickening results in brown fibres with improved quality. *Plant Biotechnology Journal* 16(10):1735–1747. https://doi.org/10.1111/pbi.12910

Yatsu, L. Y., Espelie, K. E., and Kolattukudy, P. E. 1983. Ultrastructural and chemical evidence that the cell wall of green cotton fiber is suberized. *Plant Physiology* 73(2):521–524. https://doi.org/10.1104/pp.73.2.521

Zhang, X. Z., Zhou, Y. S., and Lan, J. X. 2002. Current status of research and genetic analysis of some traits of naturally colored cotton. *Chin Agric Sci Bull* 18:92–93.

Zhao, J. and Dixon, R. A. 2009. MATE transporters facilitate vacuolar uptake of epicatechin 3'-o-glucoside for proanthocyanidin biosynthesis in *Medicago truncatula* and *Arabidopsis*. *The Plant Cell* 21(8):2323–2340. https://doi.org/10.1105/tpc.109.067819