An Overview on Advances in Cotton Genome and Regulation of Fiber Development

Devesh Kumar Mishra[1], Neha Agrawal[2], Amit Choudhary[3], Vikash Kumar Yadav[4], Vrijesh Kumar Yadav[5]

[1]Plant Molecular Biology Laboratory, CSIR-National Botanical Research Institute, Rana Pratap Marg, Lucknow, 226001, India

[2]State Forensic Science Laboratory (SFSL), Hotwar, Ranchi, Jharkhand, 835217, India

[5]ICMR-SRF (Plant Molecular Biology Lab- NBRI), India
Scientific Assistant (SFSL, Ranchi Jharkhand) India

Abstract: ***In the evolution of flowering plants major driving forces for the evolution is polyploidy and whole-genome duplication. The genus Gossypium contains 45 diploid (2n = 26) and six tetraploid (2n = 52) species. During evolution, polyploidy (allo-tetraploidy) enhanced the fiber traits in allo-tetraploids as compared to diploids, resulting in higher fiber yield and quality. The advancement of G. hirsutum genome into AD-genome (allo-tetraploid) brings by the polyploidization event between diploid progenitors of A- and D-genome. The genome size of G. raimondii and G. arboreum is 880 Mb and 1,700 Mb respectively reveals that the size of D-genome is lesser than the genome size of A-genome. Initiation and elongation are the best understood developmental stages at the transcriptional level among different stages of fiber development. However, scanty knowledge is available about the SCW stage-specific transcriptional regulation and its detailed molecular mechanism. Since specific promoters regulates transcription in fiber cells hence the knowledge of novel fiber specific regulatory elements along promoters and their interacting transcription factors might be a better tool to be used in the genetic manipulation of fiber quality parameters.***

Keywords: *Gossypium*, transcription regulation, promoter, transcription factor, cotton genome, regulatory elements

1. Introduction

The cotton fiber is a single cell structure initiated on the epidermal surface of an ovule and eventually undergoes into four developmental stages namely: initiation, elongation, secondary cell wall (SCW) synthesis and maturation (Basra and Malik 1984). A single mature cotton fiber consists of the thin outer primary cell wall (PCW), an inner thick secondary cell wall (SCW), and a central lumen. In spite of being a valuable resource for the textile industries cotton fiber is an excellent model to study the cell differentiation and development. Cotton fiber is an important raw material for the textile industry. Economy of the world and daily human life is affected by the world wide cotton production. The genus *Gossypium* contains 45 diploid (2n = 26) and six tetraploid (2n = 52) species (Hawkins et al., 2006 Grover et al., 2015). The genus, *Gossypium* is a perennial shrub belongs to malvaceae family, grown mainly for the seed hair (textile fiber). The spinnable fiber is produced by two tetraploids (*G. hirsutum* and *G. barbadense*) and two diploids *Gossypium* species (*G. herbaceum and G. arboreum*).

In the evolution of eukaryotic organism's specially in flowering plants the main driving force for the evolution is polyploidy and whole-genome duplication (Soltis et al., 2014; Hegarty et al., 2008; Otto et al., 2007; Jiao et al., 2011). During evolution, polyploidy (allo-tetraploidy) enhanced the fiber traits as compared to diploids, resulting in higher fiber yield and quality (Chen et al., 2007; Paterson et al., 2012; Li et al., 2014; Guan et al., 2014). The *G. hirsutum* is an allo-tetraploid (AD-genome) species evolved through the polyploidization event between diploid progenitors of A- and D-genome during the course of evolution (Page et al., 2013). D- genome is dominating than A- genome inferred from the origin of AD genome, transposable elements (TEs) and reduction in the genome size (AD) resulted from allo-polyploidization (Li et al., 2015) during course of evolution of *Gossypium* spp. The *G. hirsutum*, in which fiber length is approximately 25 to 34 mm, is most widely cultivated cotton overall and more than 90% of lint fiber of world's total cotton production is obtained from *G. hirsutum* (Lin et al., 2009). Thus, the *G. hirsutum* tends to become the intensive target of basic and applied research.

2. Evolutionary divergence of *Gossypium*

The phylogenetic analysis revealed *Gossypium arboreum* and *Gossypium raimondii* evolved from T. cacao (Li et al., 2014; **Figure 1)**. The differences in the size of *Gossypium arboreum* (AA; 2n = 26) and *Gossypium raimondii* (DD; 2n = 26) genome are due to the insertions of long terminal repeats (5 million years back).

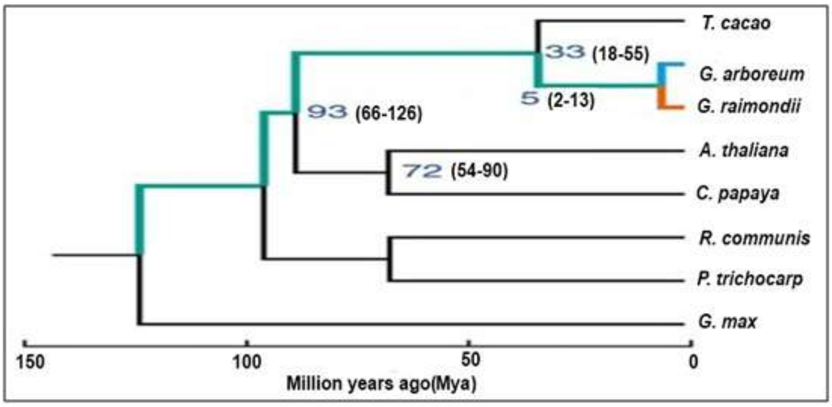

Figure 1: The schematic diagram of the *G. arborium*, *G. raimondii* and six other genomes descended from common eudicot genome ancestors. Number represent the evolutionary time in million years (Mya) (Source: Li et al., 2014).

Allo-tetraploid cotton (AADD) evolved from the natural hybridization between A and D genome species and has split into six species, including the widely cultivated *G. barbadense* (AD2) and *G. hirsutum* (AD1). The *Gossypium spp*. comprises of nine different genome groups (A, B, C, D, E, F, G and K). The genomes sizes of different genome groups (A, B, C, D, E, F, G and K) vary due to the lineage-specific proliferation of retro-transposons (Hawkins et al., 2006). The genome size of *G. raimondii* and *G. arboreum* is 880 Mb (Hendrix et al., 2005; Paterson et al., 2012; Wang et al., 2012) and approximately 1,700 Mb (Li et al., 2014) respectively implies that the genome size of D genome is less than the genome size of A genome. Genome duplication, retro-transposons and polyploidization events during course of evolutionary process in *Gossypium* spp. including flowering plants and many crop plants results in divergence of genome into different sub-genomes (Renny-Byfield et al., 2014; Gong et al., 2013). However, the molecular mechanism of occurrence of allopolyploids and the fiber traits of cotton remains limited.

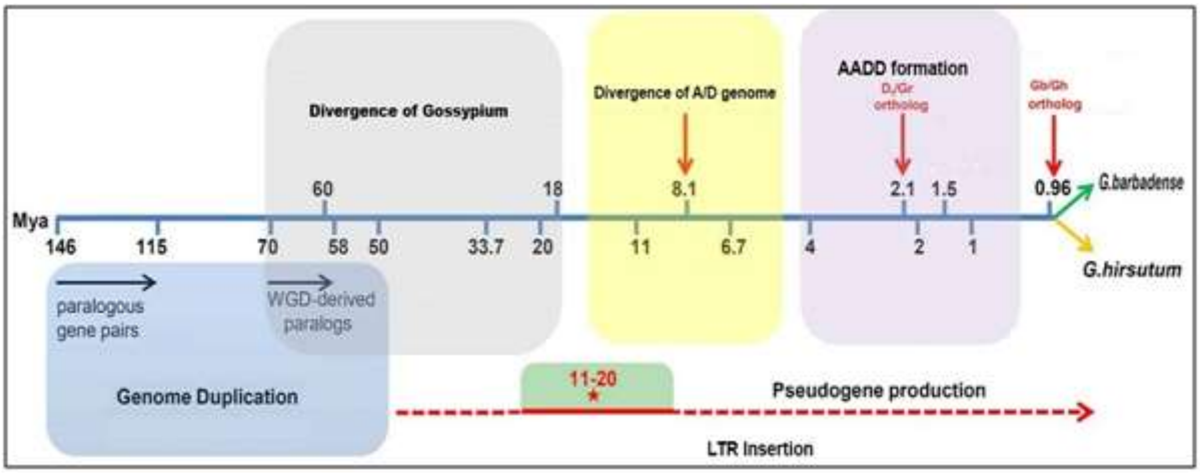

Figure 2: The schematic presentation of the evolution of allotetraploid (AADD) genome of cotton

G. hirsutum (AD1) and *G. barbadense* (AD2) splits from allotetraploid cotton (AADD) which evolved from A- and D- genome. Numbered axis represent time in million years (Mya) and evolutionary events are represented by arrows and encircled in boxes. Red star indicates pseudogene production. *Gr*: *G. raimondii; Gb*: *G. barbadense; Gh*: *G. hirsutum*, LTR: Long terminal repeats (Source: Liu et al., 2015).

3. Expression Pattern of Genes in Cotton Fiber Development

A number of reports have been generated to identify the genes related to fibre development. People have used microarray to identify the genes involved in fibre development of domesticated, mutant and poor fibre quality cotton cultivars (Li et al., 2002, Wu et al., 2007; Yang et al., 2006, Nigam et al., 2014). The differentially expressed genes (*DEGs*) during different stages of fiber development are required for fiber differentiation and development. Several genes from developing cotton fiber have been cloned and characterized (John & Crow, 1992; John & Keller, 1995; John & Keller 1996; Reinhart et al., 1996; Ma et al., 1995, 1997; Song & Allen, 1997; Smart et al., 1998; Whittaker & Triplett, 1999; Orford & Timmis, 2000; Liu et al., 2000). So far, only few of these genes are known to involved in the biosynthesis of cell wall components have been identified.

Initiation and elongation stage of fiber development is better understood. Some of the characterized initiation and elongation-specific regulator and genes include; *MYB25* (Machado et al., 2009; Wu et al., 2006), *MYB25-like* (Walford et al., 2011), *Sucrose synthase* (*Sus*) (Ruan et al., 2003), *GbPDF1* (Deng et al., 2012), *GhH2A12* (Hao et al., 2014), IAA (Zhang et al., 2011), *GhMYB2* (Wang et al., 2004), *GhHD1* (Walford et al., 2012) and *GhJAZ2* (Hu et al., 2016), *GhFAnnxA* (Zhang et al., 2016), *GhCaM7-like* (Cheng et al., 2016), *GhMYB109* (Pu et al., 2008) and *GhPK6* (Zhang and Liu, 2016) etc. During fiber development fiber cells undergo dramatic changes in gene expression that lead to massive changes in the level of proteins and metabolites important for fiber development (Mansoor & Paterson, 2012). A detailed microarray analysis of fiber development in contrasting genotypes of *G. hirsutum* suggested about the transcriptional reprogramming during fiber development (Nigam et al., 2014). Some fiber genes are reported to be constitutively expressing whereas others are reported to express during the specific stages (John & Crow, 1992; Delmer et al., 1995; John & Keller, 1995; Ma et al., 1997; Song & Allen, 1997; Smart et al., 1998; Loguerico et al., 1999; Whittaker & Triplett, 1999; Orford & Timmis, 2000; Xu et al., 2008; Nigam et al., 2014).

With the recent availability of progenitors genome sequences (Paterson et al., 2012; Wang et al., 2012; Li et al., 2014), more focus on cotton genomic studies is now shifted towards gaining insight into the mechanistic details. The study of mechanisms involved in regulation of critically important genes and metabolic networks at various levels is of immediate need to utilize the rapidly increasing cotton genomic resources for the betterment of cotton fiber.

4. Involvement of Promoter in Fiber Development

Several promoters have been reported that regulate gene expression in cotton fibers during the initiation and elongation stages (Delaney et al., 2007; Hussey et al., 2011; Larkin et al., 1996; Li et al., 2002; Li et al., 2005; Ma et al., 1995; Ni et al., 2008; Song and Allen, 1997; Wang et al., 2004; Wu et al., 2009). The promoter of *GaRDL1*, *GhTUB1* and *GhMYB25* are active early during the initiation stage (Li et al., 2002; Machado et al., 2009; Wang et al., 2004), while the glucuronosyl transferase promoter (PGhGlcAT1), has been shown to be highly active during fiber elongation (Wu et al., 2007). The fiber-specific lipid transfer protein (*FSltp4*) promoter has also been reported to be active during fiber elongation and has been shown to be functional during the synthesis of fiber cutin (Delaney et al., 2007). *GhACTIN1* is another gene expressed in fiber and its promoter has been shown to be elongation-specific (Li et al., 2002). Promoters for the genes, *GhLTP3*, *GhDET2*, *GaMYB2*, and *GhMYB109*, have also been reported to be active during initiation and some of them continue to be active until the elongation stage (Liu et al., 2000; Luo et al., 2007; Pu et al., 2008; Wang et al., 2004). There are other promoters, such as $P_{GhRING1}$, that are active in all the stages of cotton fiber development, starting from initiation to the SCW stage (Ho et al., 2010). Among different developmental stages of the fiber, initiation is the best understood at the transcriptional level. However, there are scanty knowledge about the SCW stage-specific transcriptional regulation and its detailed molecular mechanism. Recently SCW stage specific *GhGDSL* promoter have been cloned and characterized in cotton transgenic lines and showed enhanced GUS and GFP expression at SCW biosynthesis stage (19-25DPA) of fiber development. In this study through yeast one hybrid interaction (Y1H) it has been shown that *GhGDSL* promoter interact with GhMYB1 which is the regulator of SCW stage specific expression in *Gossypium hirsutum* (Yadav et al., 2017).

Efforts to understand and manipulate fiber development would be greater by an improved knowledge of the promoter elements and transcription factors that regulate fiber-specific gene expression. Promoters are DNA sequences that determine the expression level, time and tissue specificity of gene. Genetic engineering for improved agronomic traits requires specific promoters for the modification of fiber properties (John, 1998). The promoters of fiber-specific genes may regulate gene function and fiber development by altering transcription of fiber cells.

Table 1: Some conserve motif sequences of promoter

S.No.	Conserve Sequences	Motifs ID	Putative Functions	References
1	CW8G	CARGCW8GAT	A variant of CArG, binding site for AGL15	Plant Physiol (2006)
2	ACACNNG	DPBFCORE	Age & developmental regulation	Mol Cell Biol (2012)
3	GATA	GATABOX	Tissue-specific promoter activity of legumin gene	Mol Gen Genet (1989)
4	GCCGCC	GCCCORE	Light & SA-inducible expression	Plant Mol Biol (1995)
5	GRWAAW	GT1CONS	Transcriptional activation of the rbcS 3A gene	EMBO (1988)
6	GGTTAA	GT1CORE	Pathogen & salt-induced SCaM4 gene expression	Plant Physioly (2004)
7	GAAAAA	GT1GMS	Regulation of expression of histone H3 gene	Plant Cell Physioly (1989)
8	ACGTCA	HEXMOTIF	Light-regulated transcription	Plant Mol Biol (1995)
9	GATAA	IBOXCORE	Light responsive transcription of psadb gene	Plant Journal (2002)
10	TAAATGYA	L1BOXAT	Consensus sequence for plant introns	Nucleic Acids Research (1986)
11	WAACCA	*MYB*1AT	Trichome regulation, phospholipid signaling (GLABRA2) & epidermal differentiation	Plant Cell (2004)
12	YAACKG	*MYB*2CONSAT	Transcriptional activators in abscisic acid signaling	Plant Cell (2003)
13	CNGTTR	*MYB*CORE	Transcriptional activators in abscisic acid signaling	Plant Cell (2003)
14	AACGG	*MYB*CORE	Regulate flavonoid biosynthesis in stress	EMBO (1995)
15	GGATA	*MYB*ST1	Cell cycle regulation	Plant Mol Biol (2002)
16	GGATA	*MYB*ST1	Function as a transcriptional activator	EMBO (1994)
17	CATGTG	MYCATERD1	Function in dehydration stress and dark-induced senescence	Plant Cell (2004)

18	GANTTNC	EECCRCAH1	Regulate light & tissue specific expression	Plant Cell (1990)
19	CACATG	MYCATRD22	Regulation of ABA-induced transcription	Plant Cell (1997)
20	CANNTG	MYCCONSAT	Phenylpropanoid biosynthesis genes	Plant Cell (2005
21	ACTTTA	BBF1ARR	Tissue-specific expression & auxin induction	Plant Cell (1999)
22	AGAAA	POLLEN1AT52	Play role as polyadenylation signal for plant genes	Nucleic Acids Res (1987)
23	ACTCAT	PREATPRODH	Pollen-specific transcription	Plant Mol Biol (1998)
24	CATGCA	RYREPEAT	Required for phytochrome regulation	Plant Cell (1996)
25	ATGGTA	S1FBOX	ABA-responsive complex for seed-specific expression	Plant Mol Biol (1999)
26	RTTTTTR	SEF4MOTIF	Developmental regulation of a house-keeping gene	Mol Cell Biol (1993)
27	GGGCC	SORLIP2AT	Nuclear factors interact with soybean embryo factor	Plant Cell (1989)
28	TTATCC	SREATMSD	Phytochrome A regulated expression, circadian regulation	Plant Physiol (2003)
29	GAGAC	SURECOREAT	Circadian regulation	Plant Physiol (2003)
30	TAAAG	TAAAGSTKST1	Regulate gene expression in axillary bud outgrowth	Plant Physiol (2005)
31	TATTTAA	TATABOX	Regulation of a light-regulated gene	Plant Physiol (1995)
32	T-3ATATA	TATABOX	Binding site for OsTBP2 in pal promoter	Plant Cell (2002)
33	TATCCA	TATCCAOS	TATA and *MYB* binding site in GapC4 promoter	Biochem Biophys Acta (2001)
34	ATTC-3GC	TE2F2NTPCNA	Required for high-level expression of alpha-Amylase	Plant Cell Physiol (2001)
35	TGACGT	TGACGTV	Involved in transcriptional activation	Plant Journal (2002)
36	TGAC	WRKY	Phyto hormonal signalling GA & ABA	EMBO (1999)
37	TGACY	WBox	Transcription activation	Biochim Biophys Acta (2004)
38	CAAT	CAATBox	Act cooperatively with HSEs to increase promoter activity	Plant Physiol (2002)
39	GATA	GATABox	Required for light regulation and tissue specific expression	Plant cell (1990)

Several promoters have been shown to drive fiber-specific gene expression in transgenic plant (Rinehart et al., 1996, Wu et al., 2009, Wu et al., 2007, Delaney et al., 2007, Li et al., 2002, Harmer et al., 2002, Botha et al 2011, Ni et al., 2008, Pyee and Kolattukudy 1995, Wang et al., 2004). Fiber characteristics have also been manipulated by coupling fiber-specific promoters to specific transgenes that may improve fiber quality (John and Keller 1996, Rinehart et al., 1996).

5. Transcription regulation

Promoters from some of the fiber specific genes have been isolated and their activities have been evaluated in transgenic plants. The conserve *cis*-regulatory elements help in transcription of genes in fiber cells (**Table 1**). Several of these promoters are proven to be applicable in genetic engineering of fiber. Trichomes in *Arabidopsis* and tobacco show a number of structural and genetic similarities to cotton fibers (Wang et al., 2004, Humphries et al., 2005). Because promoter analysis in cotton is limited by the resources and extended timeline (minimum 12 months), the activity of cotton fiber-specific promoters has been also monitored in the heterologous model systems (Hsu et al., 1999, Liu et al., 2000, Wang et al., 2004, Wu et al., 2006).

The developing cotton fiber is an excellent model for unraveling the fundamental process of plant cell growth, differentiation, cell wall biogenesis and number of fiber mutants is available for exploring the regulatory mechanism of fiber growth and developmental processes. Recently much of work has been done for promoter characterization specific to fiber tissue in transgenic plants (**Table 2**).The promoter of *LTP6* gene was shown to drive expression of *gus* gene specifically in tobacco leaf trichome cells and may contain similar regulatory factors for tissue specific expression of gene in fiber cells (Hsu et al., 1999). Recently, the promoter of *GhTUB1* gene was shown to direct *gus* expression in transgenic cotton fiber cells (Li et al., 2002).

Table 2: Some Trichome/Fiber specific promoter, expression stages and putative function in transgenic plants

S.No.	Promoter	Expression stage	Putative function	References
1	GhLTP6	Initiation & Elongation	Fiber morphogenesis	Hsu et al., 1999
2	GhGlcAT1	Fiber elongation	Cell wall synthesis	Wu et al., 2007
3	GhACTIN1	Elongation	Fiber elongation	Li et al., 2002
4	GhE6	Elongation	Fiber growth	John & Crow 1992
5	FSltp4	Fiber elongation	Synthesis of fiber cutin	Delaney et al., 2007
6	GhLTPxGH3	Initiation & Elongation	Fiber morphogenesis	Ma et al., 1995
7	AtTSG	Initiation	Fiber morphogenesis	Ni et al., 2008
8	Gh*MYB*109	Initiation- Elongation	Fiber formation	Wang et al , 2008
9	GaRDL1	Initiation	Fiber formation	Wang et al., 2004
10	AtGL1	Initiation	Fiber morphogenesis	Larkin et al., 1993
11	FbL2A	Elongation and wall synthesis	Fiber growth	Rinehart et al., 1996
12	GhLTP3	Initiation & Elongation	Fiber morphogenesis	Liu et al., 2000
13	GhTUB1	Early stage	Fiber initiation	Li et al., 2002
14	GhGlcAT1	Elongation	Fiber development	Wu et al., 2007
15	GhTUA9	Elongation	Fiber growth & development	Li et al., 2007
16	Ga*MYB*2	Initiation & elongation	Fiber development	Wang et al., 2004
17	Gh*MYB*25	Initiation	Fiber formation and growth	Machado et al., 2009
18	GaHOX	Initiation & Elongation	Fiber formation	Guan et al., 2008

19	GhCesA4	Elongation & SCW	Cell wall morphogenesis	Wu et al., 2009
20	GhSUS3	Initiation	Fiber morphogenesis	Ruan et al., 2009
21	GhRING1	Initiation to SCW	Fiber formation	Ho et al., 2010
22	GhXTH1	Elongation	Fiber growth	Lee et al., 2010
23	GhSCFP	Initiation and Elongation	Fiber development	Hou et al., 2008
24	GhDET2	Initiation & Elongation	Fiber formation	Luo et al., 2007
25	GhGDSL	SCW	Cell wall biosynthesis	Yadav et al.,2017

The promoter of a cotton lipid transfer protein gene, *FSltp4*, was isolated and shown to direct fiber-specific transcription in cotton. In transgenic tobacco, this promoter was strongly active in foliar trichomes. Deletion analysis of the promoter identified an AT-rich 84 bp fiber specificregion (FSR) necessary for activity solely in the fiber cells. Cotton fiber proteins that bind the FSR were isolated using a yeast one-hybrid assay (Delaney et al., 2007). There are several *CesA* genes which belongs to multi-gene family have been reported for primary and secondary cell wall cellulose synthesis in *Arabidopsis* (Fagard et al., 2000). In Comparison to these cotton homologs such as *GhCesA1*, *GhCesA2*, and *GhCesA4*, were expressed at high levels in cotton fibers during secondary cell wall synthesis while *GhCesA3* gene expression was observed during both primary and secondary cell wall biogenesis stages (Kim and Triplett, 2001). *GhCesA4* promoter activity was conferred by GUS activity in transgenic tobacco plant (Liu et al., 2000; Wu et al., 2007). In the study of *GhGlcAT1* promoter (Wu et al., 2007) it was observed that it direct temporal, spatial and inducible expression of a reporter gene in transgenic tobacco and shows transcriptional regulation in transgenic plants. *Ltp6* and *Ltp3* promoter activity were conferred in transgenic tobacco (Ma et al., 1997). The transgenic plants demonstrated highly specific *gus* expression in the leaf and stem trichomes, with some expression in the vascular tissue and leaf margin. Deletion assay were performed to find out the promoter regions necessary for trichome specific gene expression (Hsu et al., 1999, Liu et al., 2000) and with removal of the basal promoter elements CAAT and TATA boxes, Ltp6 promoter resulted in a progressive reduction in GUS activity therefore basal promoter activity ceased in transgenic plant (Hsu et al., 1999). Recent experiment on RDL1 promoter provide envisage that cotton and *Arabidopsis* use similar transcription factors for regulating trichomes and GaMYB2 may be a key regulator of cotton fiber development (Wang et al 2004). Thus, promoter regulatory regions and interacting factors may be employed for the improvement of cotton fiber trait.

6. Conclusion

Cotton fiber is one of the important commercial crops worldwide and its development takes place in the stage specific manner. The regulation of cotton fiber developmental stages is very important for fiber quality parameters. Promoter is crucial region of gene which plays an important role in the fiber-specificity and regulation of timing of expression of genes. The promoters of these genes are of prime interest because they control transcription in the fiber cell and may be used in the genetic manipulation of fiber quality. Promoter allows the expression of transgenes targeted to the fibers and avoid detrimental effects on growth and phenotype elsewhere within the plant. This will help in developing new strategies to better utilize the promoter for the regulation of gene expression during fiber developmental processes and ultimately in improving the cotton fiber quality and yield.

References

[1] Basra A, & Malik CP (1984). Development of the cotton fiber. *International Review of Cytology* 89: 65-113.

[2] Botha J, Pinard D, Creux N, Hussey S, Maritz-Olivier C, Spokevicius A, Bossinger G, Mizrachi E, & Myburg A (2011). Characterising the role of the *Eucalyptus grandis* SND2 promoter in secondary cell wall biosynthesis. *BioMed Central Proceeding* 5 (7): 105).

[3] Chen ZJ, Scheffler BE, Dennis E, & Triplett BA (2007). Towards sequencing cotton (*Gossypium*) genomes. *Plant Physiology* 145: 1303-1310.

[4] Cheng Y, Lu L, Yang Z, Wu Z, Qin W, Yu D, Ren Z, Li Y, Wang L, Li F (2016). *GhCaM7-like*, a calcium sensor gene, influences cotton fiber elongation and biomass production. *Plant Physiol Biochem* 109: 128-136.

[5] Delaney SK, Orford SJ, Martin-Harris M, & Timmis JN (2007). The fiber specificity of the cotton *FSltp4* gene promoter is regulated by an AT-rich promoter region and the AT-hook transcription factor GhAT1. *Plant Cell Physiology* 48: 1426-1437.

[6] Fagard M, Desnos T, Desprez T, Goubet F, Refregier G, Mouille G, McCann M, Rayon C, Vernhettes S, & Hofte H (2000). PROCUSTE1 encodes a cellulose synthase required for normal cell elongation specifically in roots and dark-grown hypocotyls of Arabidopsis. *Plant Cell* 12: 2409-2424.

[7] Gong L, Kakrana A, Arikit S, Meyers BC, & Wendel JF (2013). Composition and expression of conserved microRNA genes in diploid cotton (Gossypium) species. *Genome biol Evol* 5: 2449-2459

[8] Guan X, Pang M, Nah G, Shi X, Ye W, Stelly DM, & Chen ZJ (2014). miR828 and miR858 regulate homoeologous MYB2 gene functions in *Arabidopsis* trichome and cotton fibre development. *Nature communications* 5.

[9] Grover CE, Zhu X, Grupp KK, Jareczek JJ, Gallagher JP, Szadkowski E, & Wendel JF (2015). Molecular confirmation of species status for the allopolyploid cotton species, Gossypium ekmanianum Wittmack. *Genetic resources and crop evolution* 62(1): 103-114.

[10] Hao J, Chen S, Tu L, Hu H, & Zhang X (2014). GhH2A12, a replication-dependent histone H2A gene from Gossypium hirsutum, is negatively involved in the development of cotton fiber cells. *Plant cell reports* 33(10): 1711-1721.

[11] Harmer, S, Orford, S, & Timmis, J. (2002) Characterisation of six α-expansin genes in *Gossypium hirsutum* (upland cotton). *Molecular Genetics & Genomics* 268(1): 1-9.

[12] Hawkins J. S, Kim H, Nason J. D, Wing R. A. & Wendel J. F. (2006) Differential lineage-specific amplification of transposable elements is responsible for genome size variation in Gossypium. Genome Res. 16, 1252–1261.

[13] Hegarty, M. J. & Hiscock, S. J. (2008) Genomic clues to the evolutionary success of polyploid plants. Curr Biol. 18, R435–444. Hu, H, He, X, Tu, L, Zhu, L, Zhu, S, Ge, Z, & Zhang, X. (2016). GhJAZ2 negatively regulates cotton fiber initiation by interacting with the R2R3-MYB transcription factor GhMYB25-like. The Plant Journal, 88(6), 921-935.

[14] Hendrix, B. & Stewart, J. M. Estimation of the nuclear DNA content of gossypium species. Ann Bot. 95, 789–797 (2005).

[15] Ho, M.H, Saha, S, Jenkins, J.N, & Ma D.P. (2010) Characterization and promoter analysis of a cotton RING-type *ubiquitin ligase* (E3) gene. *Molecular Biotechnology* 46: 140-148.

[16] Hussey, S.G, Mizrachi, E, Spokevicius, A.V, Bossinger, G, Berger, D.K, & Myburg, A.A. (2011) SND2, a NAC transcription factor gene, regulates genes involved in secondary cell wall development in Arabidopsis fibres and increases fibre cell area in Eucalyptus. *BioMed Central Plant Biology* 11: 173.

[17] Hsu, C.Y.C.R, Jenkins, J.N, Ma, D.P. (1999) Analysis of promoter activity of cotton lipid transfer protein gene LTP6 in transgenic tobacco plants. *Plant Science* 143: 63-70.

[18] Jiao, Y, Wickett, N. J, Ayyampalayam, S, Chanderbali, A. S, Landherr, L, Ralph, P. E, & Soltis, D. E. (2011). Ancestral polyploidy in seed plants and angiosperms. Nature, 473(7345), 97-100.

[19] John, M. E. (1998) Genetic engineering of cotton fiber. In Cotton (pp. 313-331). Springer Berlin Heidelberg.

[20] John, M.E & Crow, L.J. (1992) Gene expression in cotton (Gossypium hirsutum L.) fiber: cloning of the mRNAs. *Proceedings of the National Academy of Sciences* 89(13): 5769-5773.

[21] John, M.E. & Keller, G. (1995) Characterization of mRNA for a proline-rich protein of cotton fiber. *Plant Physiology* 108(2): 669-676.

[22] John, M.E, & Keller, G. (1996) Metabolic pathway engineering in cotton: biosynthesis of polyhydroxybutyrate in fiber cells. *Proceedings of the National Academy of Sciences* 93(23): 12768-12773.

[23] Kim, H. J, & Triplett, B.A. (2001) Cotton fiber growth in planta and in vitro. Models for plant cell elongation and cell wall biogenesis. *Plant Physiol* 127: 1361-1366.

[24] Kim Y J, Wang R, Gao L, Li D, Xu C, Mang H, Jeon J, Chen X, Zhong X, Kwak J M (2016). POWERDRESS and HDA9 interact and promote histone H3 deacetylation at specific genomic sites in Arabidopsis. *Proc Natl Acad Sci* 18618.

[25] Li, F, Fan, G, Lu, C, Xiao, G, Zou, C, Kohel, R. J, & Liang, X. (2015). Genome sequence of cultivated Upland cotton (Gossypium hirsutum TM-1) provides insights into genome evolution. Nature biotechnology, 33(5), 524-530.

[26] Li, F, Fan, G, Wang, K, Sun, F, Yuan, Y, Song, G, Li, Q, Ma, Z, Lu, C, Zou, C, Chen, W, Liang, X, Shang, H, Liu, W, Shi, C, Xiao, G, Gou, C, Ye, W, Xu, X, Zhang, X, Wei, H, Li, Z, Zhang, G, Wang, J, Liu, K, Kohel, R.J, Percy, R.G, Yu, J.Z, Zhu, Y.X, Wang, J, & Yu, S. (2014) Genome sequence of the cultivated cotton Gossypium arboreum. *Nature Genetics* 46(6): 567-572.

[27] Li, X.B, Cai, L, Cheng, N.H, & Liu, J.W. (2002) Molecular characterization of the cotton *GhTUB1* gene that is preferentially expressed in fiber. *Plant Physiology* 130: 666-674.

[28] Li, X.B, Fan, X.P, Wang, X.L, Cai, L, & Yang, W.C. (2005) The cotton *ACTIN1* gene is functionally expressed in fibers and participates in fiber elongation. *Plant Cell* 17: 859-875.

[29] Larkin, J.C, Young, N, Prigge, M, & Marks, M.D. (1996) The control of trichome spacing and number in *Arabidopsis*. *Development* 122: 997-1005.

[30] Liu, H.C, Creech, R.G, Jenkins, J.N, & Ma, D.P. (2000) Cloning and promoter analysis of the cotton *lipid transfer protein* gene (*Ltp3*). *Biochimica et Biophysica Acta* 1487: 106-111.

[31] Lin, Z, Zhang, Y, Zhang, X, & Guo, X. (2009). A high-density integrative linkage map for *Gossypium hirsutum*. *Euphytica* 166(1): 35-45.

[32] Liu, X, Zhao, B, Zheng, H. J, Hu, Y, Lu, G, Yang, C. Q, ... & Zhou, Y. (2015). Gossypium barbadense genome sequence provides insight into the evolution of extra-long staple fiber and specialized metabolites. Scientific reports, 5, 14139.

[33] Luo, M, Xiao, Y, Li, X, Lu, X, Deng, W, Li, D, Hou, L, Hu, M, Li, Y, & Pei, Y. (2007). *GhDET2*, a steroid 5alpha-reductase, plays an important role in cotton fiber cell initiation and elongation. *Plant Journal* 51: 419-430.

[34] Luo M, Tai R, Yu CW, Yang S, Chen CY, Lin WD, Schmidt W, & Wu K (2015). Regulation of flowering time by the *histone deacetylase (HDA5)* in *Arabidopsis*. *Plant J* 82: 925-936.

[35] Machado A, Wu Y, Yang Y, Llewellyn DJ, & Dennis ES (2009). The MYB transcription factor *GhMYB25* regulates early fibre and trichome development. *Plant J* 59: 52-62.

[36] Ma DP, Liu HC, Tan H, Creech RG, Jenkins JN, & Chang YF (1997). Cloning and characterization of a cotton lipid transfer protein gene specifically expressed in fiber cells. *Biochimica et Biophysica Acta* 1344(2): 111-114.

[37] Ma DP, Tan H, Si Y, Creech RG, & Jenkins JN (1995). Differential expression of a *lipid transfer protein* gene in cotton fiber. *Biochimica et Biophysica Acta* 1257: 81-84.

[38] Mansoor S, & Paterson AH (2012). Genomes for jeans: cotton genomics for engineering superior fiber. *Trends in Biotechnology* 30(10): 521-527.

[39] Ni Y, Wang X, Li D, Wu Y, Xu W, & Li X (2008). Novel cotton homeobox gene and its expression profiling in root development and in response to stresses and phytohormones. *Acta Biochimica et Biophysica* 40: 78-84.

[40] Nigam D, Kavita P, Tripathi RK, Ranjan A, Goel R, Asif M, Shukla A, Singh G, Rana D, & Sawant SV (2013). Transcriptome dynamics during fibre development in contrasting genotypes of *Gossypium hirsutum* L. *Plant Biotechnol Journal* 12: 204-218.

[41] Otto SP (2007). The evolutionary consequences of polyploidy. *Cell* 131: 452-462.

[42] Paterson, A. H, Wendel, J. F, Gundlach, H, Guo, H, Jenkins, J, Jin, D, & Yoo, M. J. (2012). Repeated polyploidization of Gossypium genomes and the evolution of spinnable cotton fibres. *Nature* 492(7429): 423-427.

[43] Page JT, Huynh MD, Liechty ZS, Grupp K, Stelly D, Hulse AM, Ashrafi H, Van Deynze A, Wendel JF, & Udall JA (2013). Insights into the evolution of cotton diploids and polyploids from whole-genome re-sequencing. *Genes Genomes Genetics* 3(10): 1809-1818.

[44] Pu L, Li Q, Fan X, Yang W, & Xue Y (2008). The R2R3 MYB transcription factor *GhMYB109* is required for cotton fiber development. *Genetics* 180: 811-820.

[45] Pyee J, & Kolattukudy PE (1995). The gene for the major cuticular wax-associated protein and three homologous genes from broccoli (*Brassica oleracea*) and their expression patterns. *Plant Journal* 7: 49-59.

[46] Rinehart JA, Petersen MW, & John ME (1996). Tissue-specific and developmental regulation of cotton gene *FbL2A*. Demonstration of promoter activity in transgenic plants. *Plant Physiology* 112: 1331-1341.

[47] Soltis DE, Visger CJ, & Soltis PS (2014). The polyploidy revolution then and now: Stebbins revisited. Am J Bot. 101, 1057–1078.

[48] Song, P, & Allen, R.D. (1997). Identification of a cotton fiber-specific acyl carrier protein cDNA by differential display. *Biochimica et Biophysica Acta* 1351: 305-312.

[49] Song Q, Guan X, & Chen ZJ (2015). Dynamic roles for small RNAs and DNA methylation during ovule and fiber development in allotetraploid cotton. PLoS Genet, 11(12), e1005724.

[50] Wang K, Wang Z, Li F, Ye W, Wang J, Song G, & Zou C (2012). The draft genome of a diploid cotton *Gossypium raimondii*. *Nature genetics* 44(10): 1098-1103.

[51] Renny-Byfield S, & Wendel JF (2014). Doubling down on genomes: polyploidy and crop plants. *Am J Bot* 101: 1711–1725.

[52] Wan QGX, Yang N, Wu H, Pan M, Liu B, Fang L, Yang S, Hu Y, Ye W, Zhang H, Ma P, Chen J, Wang Q, Mei G, Cai C, Yang D, Wang J, Guo W, Zhang W, Chen X, Zhang T (2016). Small interfering RNAs from bidirectional transcripts of GhMML3_A12 regulate cotton fiber development. *New Phytol* 210: 1298-1310.

[53] Wang S, Wang JW, Yu, N Li, CH, Luo B, Gou JY, Wang LJ, & Chen XY (2004). Control of plant trichome development by a cotton fiber *MYB* gene. *Plant Cell* 16: 2323-2334.

[54] Wu AM, Lv SY, & Liu JY (2007). Functional analysis of a cotton glucuronosyl transferase promoter in transgenic tobaccos. *Cell Research* 17: 174-183.

[55] Wu AM, Ling C, & Liu JY (2006). Isolation of a cotton reversibly glycosylated polypeptide (*GhRGP1*) promoter and its expression activity in transgenic tobacco. *Journal of Plant Physiology* 163(4): 426-435.

[56] Wu AM, Hu JS, & Liu JY (2009). Functional analysis of a cotton *cellulose synthase A4* gene promoter in transgenic tobacco plants. *Plant Cell Reports* 28(10): 1539-1548.

[57] Orford SJ, & Timmis JN (2000). Expression of a lipid transfer protein gene family during cotton fiber development. Biochimica et Biophysica Acta. 1483(2): 275-284.

[58] Liu HC, Creech RG, Jenkins JN, & Ma DP (2000). Cloning and promoter analysis of the cotton *lipid transfer protein* gene (*Ltp3*). *Biochimica et Biophysica Acta* 1487: 106-111.

[59] Whittaker DJ, & Triplett BA (1999). Gene-specific changes in alpha-tubulin transcript accumulation in developing cotton fibers. *Plant Physiology* 121(1): 181-188.

[60] Delmer DP (1999). Cellulose biosynthesis: exciting times for a difficult field of study. *Annual review of plant biology* 50(1): 245-276. Deng, F, Tu, L, Tan, J, Li, Y, Nie, Y, & Zhang, X, (2011) GbPDF1 is involved in cotton fiber initiation via the core cis-element HDZIP2ATATHB2. *Plant Physiology* 158: 890-904.

[61] Ruan YL, & Chourey PS (1998). A fiberless seed mutation in cotton is associated with lack of fiber cell initiation in ovule epidermis and alterations in sucrose synthase expression and carbon partitioning in developing seeds. *Plant Physiology* 118(2): 399-406.

[62] Walford SA, Wu Y, Llewellyn DJ, & Dennis ES (2011). *GhMYB25-like*: a key factor in early cotton fibre development. *Plant Journal* 65: 785-797.

[63] Loguerico LL, Zhang JQ, & Wilkins TA (1999). Differential regulation of six novel *MYB domain* genes defines two distinct expression patterns in allotetraploid cotton (*Gossypium hirsutum* L). *Molecular and General Genetics* 261(4-5): 660-671.

[64] Smart LB, Vojdani F, Maeshima M, & Wilkins TA (1998). Genes involved in osmoregulation during turgor-driven cell expansion of developing cotton fibers are differentially regulated. *Plant Physiology* 116(4): 1539-1549.

[65] Zhang B, Liu JY (2016). Cotton cytosolic pyruvate kinase *GhPK*6 participates in fast fiber elongation regulation in a ROS-mediated manner. *Planta* 244: 915-926.

[66] Yadav, V. K, Yadav, V. K, Pant, P, Singh, S. P, Maurya, R, Sable, A, & Sawant, S. V. (2017) GhMYB1 regulates SCW stage-specific expression of the GhGDSL promoter in the fibres of Gossypium hirsutum L. Plant Biotechnology Journal.

[67] Yang, S.S, Cheung, F, Lee, J.J, Ha, M, Wei, N.E, Sze, S.H, Stelly, D.M, Thaxton, P, Triplett, B, Town, C.D, & Chen, Z.J. (2006) Accumulation of genome-specific transcripts, transcription factors and phytohormonal regulators during early stages of fiber cell development in allotetraploid cotton. *The Plant Journal* 47(5): 761-775.

[68] Zhang F, Jin X, Wang L, Li S, Wu S, Cheng C, & Guo W (2016). GhFAnnxA affects fiber elongation and secondary cell wall biosynthesis associated with Ca^{2+} influx, ROS homeostasis and actin filament reorganization. *Plant physiology* 00597.

[69] Zhang M, Zheng X, Song S, Zeng Q, Hou L, Li D, & Xiao Y (2011). Spatiotemporal manipulation of auxin biosynthesis in cotton ovule epidermal cells enhances fiber yield and quality. *Nature biotechnology* 29(5): 453-458.

[70] Xu Z, Kohel R J, Song G, Cho J, Alabady M, Yu J, & Zhu Y (2008). Gene-rich islands for fiber development in the cotton genome. *Genomics* 92(3): 173-183.