Role of Transcription Factors and Genes Regulating Metabolic Pathways of Fibre Development in Cotton

Sainath K[1]*, Praveen K[3], Akshaya M[2] and B. P. Maruthi Prasad[4]

[1]*Department of Genetics and Plant Breeding, UAS, Raichur*
[2]*Department of Genetics and Plant Breeding, UAS, Dharwad*
[3]*Department of Agricultural Microbiology, UAS, Bangalore*
[4]*Department of Genetics and Plant Breeding, UAS, Bangalore*

Corresponding Author
Sainath K
Email: sainathkarbhari98@gmail.com

OPEN ACCESS

Keywords

Quantitative trait loci (QTLs), Transcription factor (TF), Basic-leucine zipper (bZIP), Helix–turn–helix (HTH)

How to cite this article:

Sainath, K., Praveen, K., Akshaya, M. and Prasad, B. P. M. 2025. Role of Transcription Factors and Genes Regulating Metabolic Pathways of Fibre Development in Cotton. *Vigyan Varta* 6 (12): 11-16.

ABSTRACT

Cotton (*Gossypium* spp.) is one of the most economically important fiber crops worldwide, providing raw material for the textile industry. Upland cotton (*G. hirsutum*) accounts for about 90% of global production. Cotton fiber, a single elongated epidermal cell derived from the seed coat, undergoes a complex developmental process consisting of initiation, elongation, secondary cell wall thickening, and maturation. This process is regulated by intricate molecular networks involving various transcription factors (TFs) and metabolic pathways. Among these, MYB and HD-ZIP transcription factors play pivotal roles in epidermal cell differentiation, trichome and fiber initiation, and secondary wall biosynthesis. Carbohydrate and fatty acid metabolism contribute to fiber cell elongation and wall formation by supplying essential substrates and energy. Additionally, quantitative trait loci (QTLs) associated with fiber quality traits such as length and strength have been identified, offering valuable targets for genetic improvement. Understanding these regulatory mechanisms provides important insights for enhancing fiber yield and quality through molecular breeding and biotechnological approaches.

INTRODUCTION

Cotton, one of the most important crops in the world, produces natural fiber materials for the textile industry. Cotton is a soft, fluffy staple fiber that grows in a boll, or protective case, around the seeds of the cotton plants of the genus *Gossypium* in the mallow family Malvaceae where cotton fiber is a specialized and elongated single epidermal cell that is derived from the seed coat. Fiber development is a delicate and complex process with cell differentiation lasting about 50 days and goes through four distinct but overlapping periods: initiation, elongation, secondary cell wall thickening, and maturation (Wendel and cronn, 2003).

The process of fibre development starts from 3 days before to 1 day post anthesis (DPA), approximately 20%–30% of the ovule epidermal cells begin to differentiate into spinnable fibers. Fiber cells then enter a rapid elongation period, with a growth rate of more than 2 mm/day up to 20 DPA. The elongation period determines the final length of fiber cells. At around 16 DPA, cellulose biosynthesis begins in large quantity and is deposited on the secondary cell wall. This period lasts until 40 DPA, followed by the dehydration and maturation of cotton fibers (Kim and Triplett, 2001; Gou *et al.,* 2007; Haigler *et al.,* 2012).

A number of factors affecting the development of cotton fibers have been identified: for example, ethylene biosynthesis plays a significant role during fiber elongation (Shi *et al.,* 2006), and very-long-chain fatty acids may be involved in cotton fiber development by activating ethylene biosynthesis (Qin *et al.,* 2007). In addition, ascorbate peroxidase also participates in cotton fiber cell development by modulating hydrogen peroxide homeostasis (Li *et al.,* 2007; Qin *et al.,* 2008). However, the regulatory mechanism of fiber development is still largely unknown.

TRANSCRIPTION FACTORS

In molecular biology, a transcription factor (TF) is a protein that controls the rate of transcription of genetic information from DNA to messenger RNA, by binding to a specific DNA sequence. Transcription factors (TFs) play essential regulatory roles by controlling the transcription rates of downstream genes during plant growth and development (Yang *et al.,* 2004). The DNA sequence that a transcription factor binds to is called a transcription factor- binding site or response element.

There are two types of transcription factors

- Basal TFs - TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH
- Sequence Specific\Regulatory TFs – DREB, MYB, NAC, WRKY etc The transcription factors have 3 domains namely,
- DNA binding domain (DBD) - which attaches to specific sequences of DNA (enhancer or promoter.
- Transcription activation domain (TAD) - which contains binding sites for other proteins such as transcription coregulators.
- Signal sensing domain (SSD) (optional) - (e.g., a ligand-binding domain), which senses external signals and, in response, transmits these signals to the rest of the transcription complex, resulting in up- or down-regulation of gene expression.

There are certain families of transcription factors based on DBD, such as, Basic-leucine zipper (bZIP), Helix-turn-helix, Basic helix-loop-helix, Homeo-domain proteins, Zinc fingers etc.

MYB Transcription Factors

- Most abundant Transcription factors in the Malvacae family.
- The Gr genome contains over 200 R2R3 MYBs
- The functions of MYB proteins have been investigated in numerous plant species such as Arabidopsis, maize, rice, petunia, snapdragon, grapevine poplar and apple.
- **MYB proteins are characterized by**;

1. highly conserved DNA-binding domain: the MYB domain.
2. consists of up to four imperfect amino acid sequence repeats (R)
3. 52 amino acids, each forming **three a–helices**
4. second and third helices of each repeat build a helix–turn–helix (HTH) Structure (Ramsay *et al.,* 2001)

STRUCTURE OF MYB DOMAIN

- 3 regularly spaced tryptophan hydrophobic core in the 3D HTH structure.
- The third helix of each repeat is the ‘‘recognition helix’’ that makes direct contact with DNA and intercalates in the major groove.

- **MYBMIXTA-like (MML)** MYB (myeloblastosis) Transcription factors form the subgroup 9 of R2R3- MYBs
- First characterized member was MIXTA from *Antirrhinum majus*.
- Contain the signature protein motif AQWESARxxAExRLxRES
- MML genes have been shown to be important regulators of:
- Epidermal cell differentiation specifying cell shape in petals,
- Vegetative trichome initiation and branching and seed fiber initiation. (Riechmann *et al.,* 2000)

Homeodomain Leucine Zippers (HD-Zip) Transcription Factors

- **HD-ZIP** family of TFs, are unique to plant kingdom consists of four sub families basically consists of leucine zippers.
- A common feature of all four sub-families is the presence of a leucine zipper domain (Zip) adjacent to HD, which is important for homo- and hetero-dimerization.
- HD-Zip I subfamily vs HD-Zip II subfamily have conserved “CPSCE” motif
- HD-Zip III and IV TFs contain a Steroidogenic Acute Regulatory protein-related lipid Transfer (START) domain and
- a conserved START-associated domain (HD-SAD), that are absent in HD-Zip I and II proteins.

HD-Zip -IV TFs consist of four **conserved domains**.

- A highly conserved HD domain 60 or 61 amino acid residues binding to a specific DNA sequence by forming a structure composed of three α-helices.
- HD-Zip IV proteins preferentially binds to an **11 bp-long palindromic** sequence 5'-GCATT(A/T) AATGC-3’,
- That partly overlaps with the sequence of the L1 box (5'-TAAATG(C/T)A-3'). The L1 box is responsible for specific gene expression in the epidermal L1 layer.
- The second conserved domain characteristic of HD-Zip IV TFs is a

leucine zipper (Zip), which is specific for the HD-Zip IV structure

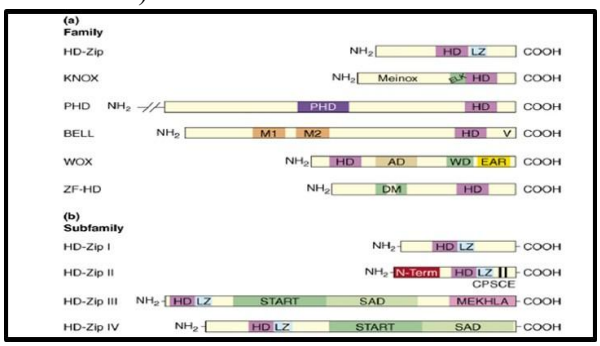
- ❑ Functionally bipartite dimerization leucine zipper-loop-zipper (ZLZ) motif.
- ❑ It is located immediately after the third helix of HD.
- ❑ The **third domain which** is a START domain, which is composed of approximately 200 amino acids residues. involved in

I. signal transduction and

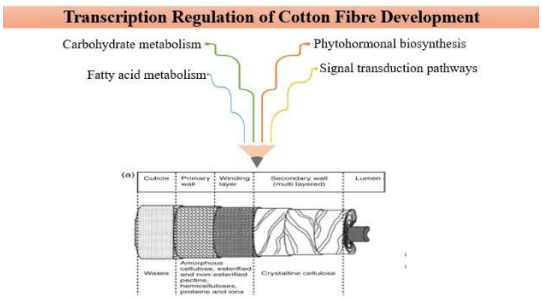
II. direct regulation of transcription by binding and

III. transporting steroid-type phytohormones and/or other lipid molecules.

IV. The binding affinity of TFs containing a START domain to specific DNA elements may be affected either by a direct protein-lipid/sterol interaction or by an interaction of lipid/sterol with a partner protein, which is bound to the same promoter region. (Ariel *et al.,* 2007)



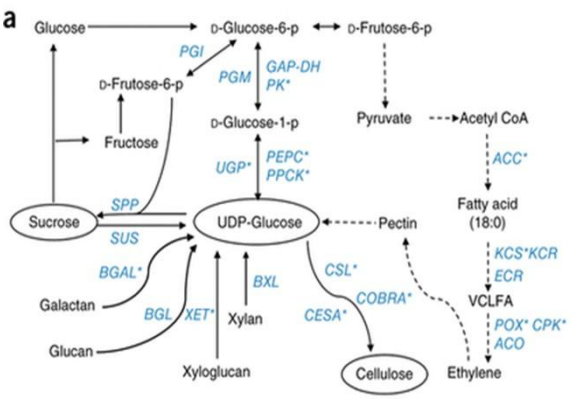
Transcription Regulation of Cotton Fibre Development



CARBOHYDRATE METABOLISM

- ❑ Sugars are the basic source of energy and carbon skeletons for all biomolecules and they are required for the regulation of cell homeostasis and synthesis of cell wall precursors.
- ❑ UDP-D-glucose (UDP-Glc) is a main metabolite and precursor for cellulose, hemi –cellulose, and and cellulose.
- ❑ **Carbohydrates** such as raffinose family oligosaccharides (RFOs) are the main storage forms of carbohydrates in the seeds, which confer desiccation tolerance.
- ❑ Trehalose is another storage carbohydrate that was shown to be involved in desiccation tolerance.

FATTY ACID METABOLISM



- ❑ Fatty acid biosynthesis is another important biochemical pathway involved in fibre cell development.
- ❑ transcripts encoding the enzymes in

I. biosynthesis of very long chain fatty acids (VLCFAs),

II. cuticular wax and

III. phospholipids were down-regulated at fibre elongation stage in the fl mutant as compared to WT

QUANTITATIVE TRAITS GOVERNING FIBER QUALITY

- The genome of upland cotton is complex and large, and the genetic background of upland cotton is narrow.
- Fiber quality traits have been proven to be negatively correlated with yield traits.
- At present, hundreds of QTL related to fiber quality and yield traits have been obtained using different mapping population.
- Some stable QTL related to yield traits were obtained
- qBS-D8-1 and qLP-D6-1, many available QTL related to fiber length and fiber strength on D3 and D11 A1, D5 and D9 & A9 (Paterson *et al*., 2012)

	Trait	QTL	Flanking marker	positio n	LOD score	P.Var(%)	Popul'n	Strategy	Refernces
1	FL	**qFl21.2**	TM76374-TM76405	109-180	5.1-1.03	2.4-4	**RIL(F2;8)**	CottonSNP80K Array	Tan *et al.*, 2018
		qFl06.1	Marker3681	35-41	4	11.1	**RIL(F2:7)**	(SLAF-seq)	Ali *et al.*, 2018
		qFl16.1	Marker18806	63,41	2.4	6.6	**RIL**		
2	FS	**Qfs-Chr01.2**	TM379-TM404	27.41	3.6-5.13	5.3-8.8	**RIL F2 (6:8)**	GWAS	Liu *et al.*, 2018
		qFChr07.2	DPL0852-DPL075	69.01	2.66-9.27	5.81-19.47	**RIL (F22;8)**		
		qFS-Chr16.3	SWU2707-DPL049	15.61	2.12-2.83	4.28-6.45	**RIL (F2 6;8)**		
3	FM	**qFm-Chr07.1**	DPL0852-DPL075	69.01	2.5-7.44	5.5-24.45		GWAS	Liu et al., 2018
		QFm-24-1	*TM69870-TM6991*	*119.333*	*3.2-5.34*	*7.9-13.5*		CottonSNP80K Array	Tan *et al.*, 2018

CONCLUSION

Cotton fiber development is a highly coordinated process governed by multiple genetic and metabolic factors. Transcription factors such as MYB and HD-ZIP families act as key regulators controlling fiber initiation, elongation and secondary wall synthesis. Carbohydrate and fatty acid metabolism provide necessary precursors and energy for fiber cell growth, while QTL mapping helps identify genomic regions associated with fiber quality traits. Integrating knowledge of transcriptional regulation, metabolic pathways and genetic markers will facilitate the development of superior cotton cultivars with improved fiber quality and yield. Continued functional genomics and molecular breeding studies will be instrumental in uncovering novel genes, including orphan genes, that contribute to the genetic improvement of cotton.

REFERENCES

Ariel FD, Manavella PA, Dezar CA, Chan RL. 2007. The true story of the HD-Zip family. *Trends in Plant Science.* 12(9): 419 – 426.

Gou JY, Wang LJ, Chen SP, Hu WL, Chen XY. 2007. Gene expression and metabolite profiles of cotton fiber development. *Plant Biotechnology Journal.* 5(2): 245–257.

Haigler CH, Singh B, Wang G, Zhang D. 2012. Genomics of cotton fiber secondary wall deposition and cellulose biogenesis. In: *Plant Fibers: Advances in Research.* Springer, New York. 1(1): 1–29.

Kim HJ, Triplett BA. 2001. Cotton fiber growth in *planta* and *in vitro*: Models for plant cell elongation and cell wall biogenesis. *Plant Physiology.* 127(4): 1361–1366.

Li XB, Fan XP, Wang XL, Cai L, Yang WC. 2007. The cotton *ACTIN1* gene is functionally expressed in fibers and participates in fiber elongation. *Plant Cell.* 17(3): 859–875.

Paterson AH, Brubaker CL, Wendel JF. 2012. Genetic improvement of cotton. In: *Crop Science: Progress and Prospects.* CABI Publishing, UK. 173–200.

Qin YM, Hu CY, Zhu YX. 2007. The ascorbate peroxidase associated with

cotton fiber development is regulated by ethylene. *Plant Physiology.* 144(1): 254–263.

Qin YM, Zhu YX. 2008. How cotton fibers elongate: A tale of linear cell-growth mode. *Current Opinion in Plant Biology.* 11(6): 636–641.

Ramsay NA, Glover BJ. 2001. MYB–bHLH–WD40 protein complex and the evolution of cellular diversity. *Trends in Plant Science.* 6(3): 106–111.

Riechmann JL, Heard J, Martin G, Reuber L, Jiang C, Keddie J, *et al*. 2000. *Arabidopsis* transcription factors: Genome-wide comparative analysis among eukaryotes. *Science.* 290(5499): 2105–2110.

Shi YH, Zhu SW, Gao WH, Zhu YX. 2006. Cotton *GhACS1* gene encoding 1-aminocyclopropane-1-carboxylic acid synthase is expressed during fiber elongation and induced by IAA. *Plant Cell Reports.* 25(10): 951–958.

Wendel JF, Cronn RC. 2003. Polyploidy and the evolutionary history of cotton. *Advances in Agronomy.* 78: 139–186.