Proteomics of Cotton Fiber Development

A. Manivannan* and Amal T. Cheeran

ICAR-Central Institute for Cotton Research, Regional Station, Coimbatore, Tamil Nadu (641 003), India

Corresponding Author

A. Manivannan

✉: mani.vannan.461@gmail.com

Abstract

Cotton is one of the major sources for natural fiber with strong commercial relevance. Now a day, there is a substantial demand for the extra-long, strong and fine cotton fibres are the major choices of modern textile industry. The most efficient way to improve fiber quality is through breeding. However, it is a big challenge for cotton breeders to develop a cultivar having good fiber quality along with higher yield because a negative genetic correlation widely exists between quality and yield. However, complete draft genome sequence of cotton is now available, but functional genomic studies remain in their infancy, as this exhibit genetic constrains like recalcitrance and complex genome. Proteomics is an established complementary tool to genomics provides a powerful tool for functional analysis of cotton fiber productivity. In this review, a special emphasis is given to cotton fiber proteomics in response to fiber biogenesis, fiber quality and fiber colour that will give a better understanding for molecular basis of acquisition of fiber developmental mechanism. Various biological and molecular pathways that are mostly carried out and controlled by proteins guide this dynamic process of cotton fibre development. An integrating proteome data with genome information from cotton will provide exciting outcome for high quality fiber and yield to attain long-term goals of cotton sustainable production.

Keywords: Cotton, Development, Fiber, Metabolism, Proteomics

Introduction

Cotton fiber is emanating as a single-celled trichomes/ hairs like growth from the ovule epidermis. As a single cell fiber having abundance of cellulose deposition on cell walls makes distinctive from other plant tissues. Although, most of the epidermal cells have become seed hairs, in which around 30% could be used for spinnable (lint fibers) and others are short (fuzz fibre). Understanding fibre development could be a paramount importance in devising a strategy for cell morphogenesis, identify specific protein, and candidate genes for improving cotton fiber quality. Elucidation of these underlying molecular mechanisms in the fibre cell, fiber growth fall into four distinct but overlapping stages; initiation, elongation, cellulose biosynthesis, and maturation. Entire fibre developmental stages will take nearly 50 days, which ultimately determines the fiber quality. Initiation step starts with the differentiation of selected epidermal cells into fiber initials, which occurs near the day of flower opening (anthesis) for lint fibres (0 to 2 days); whereas, the fuzz fibers initiate growth between (5 to 10 DPA) (Stewart, 1975). Therefore, fibre cell progression is described relative to the number of DPA (Days Post Anthesis). Morphologically distinct cellsinvolve cell wall loosening due to increased turgor pressure within the central vacuole, which develops asynchronous elongation of fiber initials (3 to 25 DPA). This could be a fiber elongation upto 6 cm or may reach one-third the height of an Arabidopsis plant (Kim and Triplett, 2001). At the meantime, cell wall thickening intermittently takes place a thick secondary cell wall composed of 94% pure cellulose. Finally, fiber cells matured before the boll opens to reveal the white soft fiber within the fruit. Cotton bolls produced early in the summer will be distinct fiber structure developed on the same plant in later season.

Proteomics Methodologies

Every cell within an organism has the same genome, but the gene expression level is varying for different types of cells based on its protein function. In between genes and proteins, there is a stability gradient: gene > mRNA > protein > protein modification. Gene expression is often interpreted in terms of protein levels, but higher or lower level of gene expression does not mean that the mRNA and proteins level also be same in a specific cell. This could be a single gene having multiple or distinct proteins due to transcriptional regulation, splicing, and post-translational modifications (PTM). It brings out the limitation of genomics and highlights the importance of proteomics in crop improvement. Unlike the genome remains pretty much static, Proteome can give a much more fleshed out view of the workings of a cell and help to elucidate the protein-protein interaction. Proteome of a cell dynamically respond to the effect of environmental stimuli and hormonechanges in plants. This dynamic role regulated by PTM which will further add to the inherent complex nature of proteome. In addition, proteomics gains importance as it represents the functional status of a cell, tissue, organelle or organism at a given time.

Recent advances in protein separation by two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) followed by mass spectrometry (MS) techniques have made deeper proteome analysis by generating robust reproducible datasets. These datasets can be qualitative or quantitative according to the methods adopted, in which, quantitative analysis grouped into two broad aspects, label-based and label-free profiling. In the label-based profiling, digested peptides are labelled using stable isotopesin the MS-based proteomics. These labelled isotopes can also be incorporated into 'shotgun' type experiments like isobaric tag for relative and absolute quantitation (iTRAQ), tandem mass tag (TMT), isotope-coded affinity tags (ICAT), and stable isotope labelling with amino acids in cell culture (SILAC). Among them, iTRAQ method is now feasible in 4-plex and 8-plex isobaric tags, which has been widely used in crop proteomics. Most label-based quantification approaches have the advantage of more accurate quantitation of multiple proteins in a single experiment. However, complex samples having low concentration and incomplete labelling much more difficult to analyse. In order to overcome this issue, label-free proteomics approaches have the advantage of more accurate quantitates both relative and absolute protein quantity by utilizing signal intensity and spectral counting of peptides. It provides a cost-effective alternative to label-based approaches. Recently, targeted proteomic approaches like multiple-reaction monitoring (MRM) and parallel-reaction monitoring (PRM) can accurately quantify the modified forms of the proteins by targeting its specific peptide sequences in a complex biological samples. Therefore, a combined workflow encompassing both discovery and validation phase of quantitative proteomics techniques would be extremely beneficial for proteomics. Despite the advent of many robust high throughput proteomic tools, suitability of 2-DE coupled with MS especially cotton proteomics remains unchallenged till date, as evidenced by the number of research papers employing this technique to understand various aspects of cotton development.

Protein Extraction from Fiber Tissues

Over the past decades, there is a paradigm shift in comprehensive protein extraction protocols from various crop tissues. Protein extraction serves as the first step in the proteomics study, which relay on protein quality and quantity. Therefore, high quality protein extraction is foremost stage and plays an important role in proteomics, which could universally deploy to various crop tissues. In cotton proteomics remained to be a challenge, since the extraction of proteins from tissue was complicated by the extreme levels of interfering substances such as cell walls, phenolic compounds, oxidative enzymes and carbohydrates particularly, sucrose. Among the tissues, fiber has recalcitrance due to relatively abundant endogenous levels of cotton metabolites (Wan and Wilkins, 1994). Since 1980's, cotton fiber protein extraction began for analysing fibre development using 2D-PAGE. They were initially homogenized the fibre tissue in an aqueous buffer and precipitated the proteins by organic solvents. But, the extracted proteins on 2D-PAGE has less protein spots smeared as streaked lines due to incomplete removal of interfering compounds (Graves and Stewart, 1988). Despite its quality of protein extraction, an alternative procedure was developed using phenol extraction followed by precipitation with methanol and ammonium acetate. This protocol has been successfully used in the large-scale protein purification especially on fiber tissue (Turley and Ferguson, 1996). However, most of the phenol-based protocols not use any detergents (SDS and CTAB) in the extraction buffer. This may also cause the amount of protein from the extracted sample not enough for large-scale proteomic analysis. A modified phenol based extraction protocol using PVPP, which act as a removal of polyphenolic compounds followed by subsequent washes with 80% acetone for the removal of lipids and salts. This protocol gave satisfactory and reproductive results in 2-D protein profiles of higher resolution of protein spot and its feasibility in fiber proteomic study (Yao *et al.*, 2006). In addition, a comparative protein extraction study showed that both TCA/ acetone and phenol-based methods were effective in purifying large amounts of protein by removing interfering secondary compounds in recalcitrant tissues for Universal proteomic analysis (Saravanan and Rose, 2004). Taken together, a protein extraction protocol having phenol-based or TCA/ acetone methods followed by other detergents/ chaotropic substances in the extraction buffer that can be easily applied for efficient high quality proteins for cotton fiber proteomic analysis.

Proteomics in Fiber Initiation and Elongation

Proteomics based research on understanding the fiber initiation mechanisms of cotton was identified a number of genes and proteins up regulated in developing fibers. Cotton fiber initiation determines the ultimate number of

fibers per ovule, thereby determining fiber yield. An iTRAQ-based proteomics of ovules from the upland cotton species *Gossipium hirsutum* and its fuzzless-lintless mutant was performed. Expression of proteins involved in carboxylic acid metabolism, small-molecule metabolic processes, hormone regulation, and lipid metabolism was significantly enhanced in wild-typeovules (Wang *et al.*, 2015). Sucrose synthase activity was found to be adverse effects on fiber cell initiation a fiberless seed mutant; whereas, it accumulated in the basal areas of initiating fiber cells in wild-type plants (Ruan and Chourey, 1998). These lower translational efficiency and protein degradation occurs in the transition from sucrose synthesis to maturity stage, which is consistent with the view that the lower protein content is detected in the late development stage of cotton fibers. Understanding the exact nature of this mutation would be a crucial step toward elucidating the genetic factors that regulate profiber cell differentiation and initiation. During fiber initiation and elongation, there are some proteins establish a complex cellular network by protein modification during translation which may involve in hormone signalling *via* jasmonate acid (JA), ethylene (ET), gibberellic acid (GA), brassinosteroid (BR) and abscisic acid (ABA) pathways. In addition, Ethylene has proved to play a major role in promoting cotton fiber elongation by increased expression of 1-Aminocyclopropane-1-Carboxylic Acid Oxidase (ACO) is responsible for increased ethylene biosynthesis (Zhu *et al.*, 2021). Thus, identification of these proteins has made a substantial contribution to understanding the molecular basis of fiber development (Pang *et al.*, 2010; Wang *et al.*, 2012). To explore the mechanisms controlling fiber development, cotton fiber protein expression differences at adjacent periods identified using iTRAQ. A total of 797 distinct proteins were identified and 48 differential expressed proteins during fiber developmental transition. There are 27 upregulated proteins were identified during initial period of fiber development, which mainly involved in carbohydrate metabolism, oxidation-reduction, cytoskeleton organization, Ca^{2+} signalling *etc*. Whereas, there are 21 downregulated proteins involved in ribosome biogenesis, secondary metabolism, and signal transduction (Zhou *et al.*, 2019). A comparative proteomic analyses of early stages of fiber elongation process in Asian cotton ovules using 2-DE followed by MALDI-TOF/TOF, there are a total of 55 protein spots were found having different abundance ranging from 1 to 9 days post-anthesis (DPA). Further, these spots were identified using high-resolution mass spectrometry and gene ontology analyses. These differential proteome expressions mainly involved in the carbohydrate/ energy metabolism and redox homeostasis (Zhang *et al.*, 2016).

Proteomics in Fiber Quality and Strength

Fiber quality is one of the most important parameter that includes fiber length, uniformity, strength, and micronaire value that are linked together. Most of these parameters are determined by genetic and environmental factors that which controls the size and shape of the matured fiber. Genetic improvement in developing quality fiber is one of the main

challenges facing the cotton breeders for the last decade. Several genes and protein function that which classified as complex quantitative traits control fiber quality traits. In order to develop elite cotton germplasm with improved fiber quality, it is essential to identify the genetic factors that control fiber development and therefore directly influence fiber quality. During transition phase of cell elongation to secondary cell wall formation, cellulose molecules was highly polymerized and rearranged into acrystalline cellulose microfibrils is highly correlated with the formation of fiber quality (Lee *et al.*, 2015). In addition, high-quality fibres and common fibres both began with a logarithmic elongation and axialization of microfibrils, but the former experienced a longer duration of elongation and a faster rate of axial arrangement of microfibrils. Therefore, differential protein expression profiling during fiber development gives an idea about the molecular mechanism of the formation of super-quality cotton fiber. Most of the proteins related to fiber quality based on the cellular stress and reactive oxygen species (ROS) homeostasis that play important roles in the lignin biosynthesis by covalently bind to protein and carbohydrate cell wall molecule that which increases the tensile strength of cell wall. Therefore, high-quality fibers are associated with high expression of the proteins related to ROS homeostasis, the continuously elevated expression of ethylene synthesis *ACO* gene (Jiang *et al.*, 2022).

Proteomics in Fiber Pigment

With the advent of cotton Industrial revolution, there has been an infrequent resurgence on high quality coloured fibres for considering environment passion. The future of coloured cotton fibres was less promising than that of white cotton fibres since they tended to be short staple and low strength. Due to shortage in demand, the coloured cotton varieties were remained poor yield and thus were found unfit and unprofitable. Whereas, white fibre cotton were in high demand because of the myriad opportunities to create fabrics in various colours and patterns using chemical dyes. However, combination of short staple coloured cotton fibres with long staple stronger white fibres have been used in many parts of the world. Therefore, it is necessary to investigate the molecular mechanisms of naturally coloured cotton fiber which is responsible for the diversity of pigments especially flavonoids. A comparative proteomic analysis between green *vs.* cotton fiber using 2-DE and MALDI-TOF/TOF analysis. Interestingly, a key enzyme in the lignans biosynthesis known as phenylocumaran benzylic ether redutase-like protein (PCBER) was highly expressed in green coloured fiber compared to white fiber. It indicates that both PCBER and lignans were responsible for the discrimination of green colour among the fiber (Li *et al.*, 2018). Similarly, protein profile comparison of brown colour fiber and white color fiber using 2-DE and MALDI-TOF/TOF techniques. These results suggested that a set of complex proteins associated with pigment biosynthesis, which are mainly involved in the flavonoid biosynthesis and proanthocyanidin content was significantly higher in brown coloured fiber compared to white colour fiber (Li *et al.*, 2013).

Conclusion

On a comprehensive view point, cotton proteomics have crossed various stages of fiber developmental since its inception during the late 1980's, which encompassed application of 2-DE and MS based shotgun method to identify fiber responsive proteins involved initiation and elongation. Although, there are many low abundant differentially expressed proteins during fiber development remain obscure under the current poteomics method. Therefore, identification and purification of specific protein from cellular organelle for gaining access to the "targeted proteomics" is a necessity in pace that which successfully applied in the cotton breeding. On a future perspective, cotton proteomics, if utilized to its full potential, could make a substantial contribution in terms of cotton yield and high fiber quality, which would ultimately have a revolutionary impact on the global economy.

References

Graves, D.A., Stewart, J.M., 1988. Chronology of the differentiation of cotton (*Gossypium hirsutum* L.) fiber cells. *Planta* 175(2), 254-258.

Jiang, J., Shi, Z., Ma, F., Liu, K., 2022. Identification of key proteins related to high-quality fiber in upland cotton via proteomics analysis. *Plant Cell Reports* 41(4), 893-904.

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 Physiology* 127(4), 1361-1366.

Lee, C.M., Kafle, K., Belias, D.W., Park, Y.B., Glick, R.E., Haigler, C.H., Kim, S.H., 2015. Comprehensive analysis of cellulose content, crystallinity, and lateral packing in *Gossypium hirsutum* and *Gossypium barbadense* cotton fibers using sum frequency generation, infrared and Raman spectroscopy, and X-ray diffraction. *Cellulose* 22(2), 971-989.

Li, Y.J., Sun, S.C., Zhang, X.Y., Wang, X.F., Liu, Y.C., Fei, X., Jie, S.U.N., 2018. New clues concerning pigment biosynthesis in green colored fiber provided by proteomics-based analysis. *Journal of Integrative Agriculture* 17(1), 46-53.

Li, Y.J., Zhang, X.Y., Wang, F.X., Yang, C.L., Liu, F., Xia, G.X., Sun, J., 2013. A comparative proteomic analysis provides insights into pigment biosynthesis in brown color fiber. *Journal of Proteomics* 78, 374-388.

Pang, C.Y., Wang, H., Pang, Y., Xu, C., Jiao, Y., Qin, Y.M., Western, T.L., Yu, S.X., Zhu, Y.X., 2010. Comparative proteomics indicates that biosynthesis of pectic precursors is important for cotton fiber and Arabidopsis root hair elongation. *Molecular & Cellular Proteomics* 9(9), 2019-2033. DOI: 10.1074/mcp.M110.000349.

Ruan, Y.L., Chourey, P.S., 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.

Saravanan, R.S., Rose, J.K., 2004. A critical evaluation of sample extraction techniques for enhanced proteomic analysis of recalcitrant plant tissues. *Proteomics* 4(9), 2522-2532.

Stewart, J.M., 1975. Fiber initiation on the cotton ovule (*Gossypium hirsutum*). *American Journal of Botany* 62(7), 723-730.

Turley, R.B., Ferguson, D.L., 1996. Changes of ovule proteins during early fiber development in a normal and a fiberless line of cotton (*Gossypium hirsutum* L.). *Journal of Plant Physiology* 149(6), 695-702.

Wan, C.Y., Wilkins, T.A., 1994. A modified hot borate method significantly enhances the yield of high-quality RNA from cotton (*Gossypium hirsutum* L.). *Analytical Biochemistry* 223(1), 7-12.

Wang, X.C., Li, Q., Jin, X., Xiao, G.H., Liu, G.J., Liu, N.J., Qin, Y.M., 2015. Quantitative proteomics and transcriptomics reveal key metabolic processes associated with cotton fiber initiation. *Journal of Proteomics* 114, 16-27.

Wang, Y., Zheng, M., Gao, X., Zhou, Z., 2012. Protein differential expression in the elongating cotton (*Gossypium hirsutum* L.) fiber under nitrogen stress. *Science China Life Sciences* 55(11), 984-992.

Yao, Y., Yang, Y.W., Liu, J.Y., 2006. An efficient protein preparation for proteomic analysis of developing cotton fibers by 2-DE. *Electrophoresis* 27(22), 4559-4569.

Zhang, B., Du, S.J., Hu, J., Miao, D., Liu, J.Y., 2016. Comparative proteomic analyses of Asian cotton ovules with attached fibers in the early stages of fiber elongation process. *Proteome Science* 14(1), 1-8.

Zhou, X., Hu, W., Li, B., Yang, Y., Zhang, Y., Thow, K., Fan, L., Qu, Y., 2019. Proteomic profiling of cotton fiber developmental transition from cell elongation to secondary wall deposition. *Acta Biochim Biophys Sin (Shanghai)* 51(11), 1168-1177. DOI: 10.1093/abbs/gmz111.

Zhu, L., Dou, L., Shang, H., Li, H., Yu, J., Xiao, G., 2021. GhPIPLC2D promotes cotton fiber elongation by enhancing ethylene biosynthesis. *iScience* 24(3), 102199.