Plant Omics Journal POJ

POJ 4(6):278-287 (2011) ISSN:1836-3644

Invited Review Article

Cotton Omics in China

Xiangdong Chen, Wangzhen Guo, Tianzhen Zhang*

National Key Laboratory of Crop Genetics and Germplasm Enhancement, Cotton Research Institute, Nanjing Agricultural University, Nanjing 210095, China

***Corresponding author: cotton@njau.edu.cn**

Abstract

In the 21st century, advent of the omics era provides scientists with greater opportunities to dissect molecular mechanisms of cotton fiber development. Cotton contributes natural fiber for the worldwide textile industry; therefore, dissecting its biological properties is a very important scientific objective. Current Chinese scientists have made significant contributions to cotton omics, focusing on genomics, transcriptomics, proteomics, and metabolomics studies. Here, we review current applications to various omics in cotton, as well as future perspectives.

Keywords: cotton; fiber; omics; China.

Introduction

Cotton (*Gossypium* spp.) is one of the most important natural fiber and edible oil crops in the world. Upland cotton (*G. hirsutum* L.), with its high yield properties, accounts for about 95% of the annual worldwide cotton production; the extra-long staple (ELS) or Pima cotton (*G. barbadense* L.), which has superior quality fiber properties, accounts for the other approximately 5%. Cotton is the leading economic crop in China. The Chinese cotton research community, which includes universities, the Chinese Academy of Agricultural Sciences (CAAS) and Chinese Academy of Sciences (CAS), has made considerable progress through common efforts. These efforts are not only reflected in a large number of original publications, but Chinese scientists are having a growing influence on the international research community. Indeed, among the 2,443 research articles concentrating on *Gossypium* sciences listed in the ISI Web of Knowledge accessed database over the last 5 years (2006-2010), 482 were contributed by Chinese scientists (Table 1). We have made considerable progress in structural genomics, such as enhancement of genetic maps, mapping of important economic traits or genes, and molecular-assisted pyramid breeding. Whole-genome sequencing of Upland cotton (*G. hirsutum*) is currently being considered by the Cotton Research Institute, the CAAS in combination with the Southern Plains Agricultural Research Center, the United States Department of Agriculture (USDA). Compared to genetic approaches, "omics" involve relatively new technologies for cotton in functional genomics research. In the present review, we focus on the major advances in cotton genomics, transcriptomics, proteomics, and metabolomics in the recent years, and discuss future prospects for Chinese cotton "omics" research.

Cotton genetic and physical maps

Genome research has been demonstrated great promise for continued and enhanced genetic improvement of crop plants (Zhang et al., 2008). Here, we summarize the major recent advances in cotton structural genomic research, such as genetic and physical maps. The development of a large number of ESTs (expression sequence tags) has provided a good source of polymerase chain reaction (PCR)-based primers for targeting simple sequence repeats (SSRs). Three molecular linkage maps from interspecific hybrid (*G. hirsutum* × *G. barbadense*) populations have been reported in China. We initially developed a large number of EST-SSR markers and constructed a high-density and gene-rich genetic map containing 2247 loci and covering 3540.4 cM, with an average inter-marker distance of 1.58 cM based on the BC_1(TM-1 × Hai7124) population (Han et al., 2004, 2006; Song et al.,2005b; Guo et al., 2007, 2008). The map will provide new insights and spur future investigations of functional and evolutionary genomics, especially those associated with cotton fiber improvement. The other two genetic maps were developed by incorporating different classes of markers at Huazhong Agricultural University (Lin et al., 2003, 2005; He et al., 2007) and at the Cotton Research Institute, the CAAS (Yu et al., 2007). Recently, an integrative linkage map was reported for *G. hirsutum*, with 506 loci covering 3070.2 cM and a mean density of 6.5 cM per locus (Lin et al., 2009). In addition, we constructed the first A-genome diploid cotton intraspecific genetic linkage map consisting of 267 loci with the total length of 2508.7 cM; this revealed that A-genome chromosomes are largely collinear with A- and D-subgenome chromosomes (Ma et al., 2008a). It is imperative to construct physical maps based on bacterial artificial chromosome (BAC) for genomics research, and advances in molecular cytogenetic techniques will speed up this objective. Fluorescence *in situ* hybridization (FISH) using BAC clones as probes has commonly been applied to chromosome identification (Wang et al., 2006b; Wang et al., 2007c). Based on two BAC libraries of 0-613-2R and TM-1 (Yin et al., 2006; Hu et al., 2009), we assigned six linkage groups (LGs) A01, A02, A03, D02, D03 and D08 to chromo-

Table 1. Ranking of countries by number of articles published in ISI Web of Knowledge accessed database in last 5 years (2006-2010) with *Gossypium* as the main subject.

Country	No. of publications
USA	913
China	482
India	260
Brazil	185
Pakistan	113
Australia	92
Israel	79
Egypt	76
Uzbekistan	41
France	34
Belgium	32

somes 13, 8, 11, 21, 24, and 19 using BAC-FISH and translocations, identified all 26 chromosome-pairs in tetraploid cotton (Wang et al., 2007c), and established 13 homeologous chromosome pairs using a new chromosome nomenclature (A1-13 and D1-13) (Wang et al., 2006b). This set of BAC markers enables us to make associations between chromosomes and their genetic linkage groups, and also provides convenient and reliable landmarks for establishing physical linkage with unknown targeted sequences. Using the same method, Wang et al. (2007b) detected the homoeologous (duplicated) segments in allotetraploid cotton, which can facilitate research in genome duplications and evolutionary genomics. It also will enable us to identify all 13 *G. arboreum* chromosomes simultaneously designated as A1-A13 through standard karyotyping using multiple BAC-FISH analyses (Wang et al., 2008b). A pachytene FISH protocol with higher axial-resolution and sensitivity has been developed (Wang et al., 2009b), and used to integrate cytogenetic and linkage maps of homoeologous chromosomes A12 and D12 in allotetraploid cotton (Wang et al., 2010b). Considerable variation in genome organization, structure, and size between A12 and D12 homoeologous chromosomes was observed. We found that the distal regions of these chromosomes displayed relatively lower levels of structural and size variation than did other chromosomal segments. The highest level of variation was found in the pericentric regions in the long arms of the two homoeologs. The overall size difference between the A and D sub-genomes mainly was associated with uneven expansion or contraction among different regions of homoeologous chromosome pairs. As an initial investigation of the fate of homoeologous chromosomes resulting from polyploidy, these results have broad general implications for future sequencing and understanding of complex genomes in plant species.

Gene tagging and QTL mapping

Molecular linkage map construction has contributed greatly to our understanding of the evolution and organization of cotton genomes, but its primary purpose is to provide a common point of reference for locating specific genes and QTLs for quantitative traits (Zhang et al., 2008). The cytoplasmic male sterile restore fertility gene (Rf_1) was the first to be tagged in *Gossypium* (Guo et al., 1997), and its location was further refined using a high-resolution genetic map containing 13 markers within a genetic distance of 0.9 cM. This delimited the possible location of the Rf_1 gene to a minimum interval of approximately 100 kb spanning two clones designated 081-05K and 052-01N (Liu et al., 2003, Yin et al., 2006). These markers closely linked to the Rf_1 genes were used in MAS breeding to develop new restorer lines in our laboratory. In addition, more than 20 qualitative genes including Gl_2^e (Dong et al., 2007), ob_1ob_2 (Qian et al., 2009), Le_3Le_4 (Song et al., 2009), ms_5, ms_6

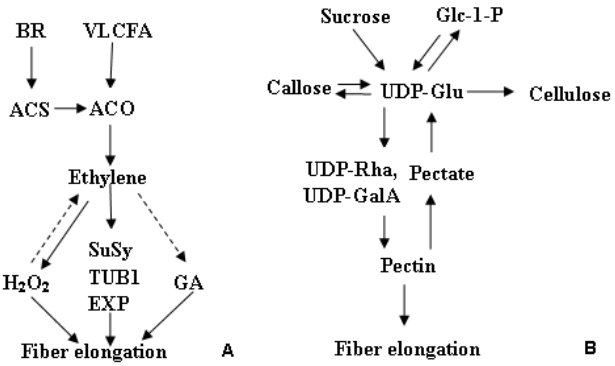

Fig 1. Regulation mechanism of cotton fiber elongation. (**A**) The signaling pathways of phytohormones. BR, brassinosteroid; VLCFAs, very-long-chain fatty acids; GA, Gibberellin. Solid lines have been confirmed by experiments; broken lines have no confirmation. Reference: Shi et al.,2006; Qin et al.,2007b; Mei et al., 2010; (**B**) The pectin biosynthesis network. Glc-1-P, Glucose 1-phosphate; UDP-GalA, UDP-D-galacturonic acid; UDP-Rha, UDP-L-rhamnose. References: Xu et al., 2007; Gou et al., 2007; Pang et al., 2010.

and ms_{15} (Chen et al., 2009), R_1 (Zhao et al., 2009), N_1 and n_2 (Song et al., 2010) and *Fw* (Wang et al., 2009a) have been tagged in this laboratory (Table 2). Important QTLs related to cotton productivity and properties, including fiber quality, overall yield and related characteristics, and resistance to diseases and insects, have been tagged (Table 2). We have uncovered an unequal distribution of QTLs between the A-subgenome (hereafter At) and D-subgenome (hereafter Dt). QTLs for fiber quality and yield more often map to Dt than At intraspecific mapping populations (Shen et al., 2005, 2006, 2007; Wang et al., 2006c, 2007d; Qin et al., 2008b); however, using interspecific mapping populations, more QTLs for fiber quality, particularly for fiber length and strength, were detected on At than Dt chromosomes (He et al., 2007; 2008b). Zhang et al. (2003) detected eight molecular markers linked with a major FS QTL ($QTLFS_1$) that explained more than 30% of the phenotypic variation in a *G. anomalum* introgression line 7235. This major QTL was stable in comparative mapping of RIL and F_2 populations (Shen et al., 2005, 2006, 2007) and was efficiently used in MAS breeding to improve fiber strength (Guo et al., 2003). In order to fine-map this QTL, three overlapping RILs, developed from a cross between 7235 and TM-1, were backcrossed to TM-1 to produce three large mapping populations. Surprisingly five tightly linked and/or clustered QTLs were detected that overlapped our previously identified major QTL region (Chen et al., 2009b). These five QTLs act like a major QTL, perhaps representing a single major gene for fiber strength, explaining a total phenotypic variance of 28.8% ~ 59.6%. A similar QTL cluster also was discovered for Verticillium wilt resistance on D7 (Wang et al., 2008a; Jiang et al., 2009). Using three elite fiber lines of Upland cotton, three pairs of homoeologous QTLs were detected (Shen et al., 2005). Similar homoeologous QTLs for FS on A8/D8 also were reported (Zhang et al., 2005; He et al., 2007). Most important QTLs for different traits have been found clustered in the same interval or in neighboring intervals. For example, Wang et al. (2006c, 2007d) tagged a stable fiber length QTL on D2; however, they simultaneously detected five significant QTLs for fiber strength, micronaire, reflectance, yellowness and maturity in four environments in Xiangzamian2

Table 2. Progress of gene tagging and QTL mapping for important traits in cotton.

Traits/genes	Parental materials	References
Rf_1 fertility-restoring gene[1]	(Zhongmiansuo12 A-1 × 0-613-2R) F_2	Guo et al., 1997
Rf_1 fertility-restoring gene[1]	CMS and the restoring lines	Liu et al., 2003
Rf_1 fertility-restoring gene[1]	XiangyuanA, ZMS12A and Sumian 16A × 0-613-2R	Yin et al., 2006
Glandless gene (*Gl^e_2*)[1]	(TM-1 × Hai1) F_2 *	Dong et al., 2007
Red plant gene (*R_1*)[1]	(Sub 16 × T586) F_2*	Zhao et al., 2009
Hybrid lethality genes(*Le_3Le_4*)[1]	(TM-1,N_1FLM and n_2 FLM × Coastland R4-4)F_2 /BC_1*	Song et al., 2009b
Open-bud duplicate genes (*ob_1ob_2*)[1]	(TM-1 × Hai7124)F_2* and (Sub18 × Hai7124 and 3–79) F_2	Qian et al., 2009
Male-sterile genes (*ms_5*,*ms_6*and *ms_{15}*)[1]	(Lang-A and Zhongkang-A × Hai7124) F_2 /BC_1*	Chen et al., 2009a
Fusarium wilt resistance gene (*Fw*)[1]	Zhongmiansuo 35 × Junmian 1	Wang et al., 2009a
Fuzzless genes (*N_1*and *n_2*)[1]	(*N_1*/*n_2* FLM × TM-1, Hai7124, Xinhai 7 and Junhai 1) F_2 /BC_1*	Song et al., 2010
Fiber strength[1]	(7235 × TM-1) F_2 and F_3	Zhang et al.,2003
Fiber strength[1]	(7235 × TM-1) F_2	Guo et al., 2003
Fiber quality[1]	7235, HS427-10, PD6992 and TM-1 × SM3	Shen et al., 2005
Fiber quality and yield[1]	(7235 × TM-1) RILs	Shen et al., 2006
Fiber quality and yield[1]	(7235 × TM-1) RILs	Shen et al., 2007
Fiber strength[1]	(7TR-133, 7TR-132, and 7TR-214 × TM-1) F_2 and $F_{2:3}$	Chen et al., 2009b
Fiber quality and yield[1]	(Simian 3 × Sumian 12) × (Zhong4133 × 8891)	Qin et al., 2008
Fiber qualities and yields [1]	Jianglingzhongmian × Zhejiangxiaoshanlushu	Ma et al., 2008
Leaf morphology and chlorophyll content[1]	(TM-1 × Hai7124) BC_1*	Song et al., 2005a
Seed physical and nutrient traits[1]	(TM-1 × Hai7124) BC_1*	Song et al., 2007
Plant architectural traits[1]	(TM-1 × Hai7124) BC_1*	Song et al., 2009c
Plant architecture traits[1]	(Zhongmiansuo 12 × J8891) RILs	Wang et al., 2006a
Fiber quality traits[1]	(Zhongmiansuo 12 × J8891) RILs	Wang et al.,2006c
Yield and yield-component traits[1]	(Zhongmiansuo 12 × J8891) RILs	Wang et al., 2007a
Fiber quality[1]	(Zhongmiansuo 12 × J8891) RILs	Wang et al., 2007d
Resistance to Verticillium wilt[1]	(Hai7124 × Junmian 1) F_2 and BC_1*	Yang et al., 2008a
Resistance to Verticillium wilt[1]	(60182 × Junmian 1) F_2	Jiang et al., 2009
Fiber quality [2]	Handan 208 × Pima90*	Lin et al., 2005
Fiber yield[2]	Handan 208 × Pima90*	He et al., 2005
Fiber quality and yield[2]	(Handan 208 × Pima 90) $F_{2:3}$ *	He et al., 2007
Fiber quality[2]	Handan 208 × Pima 90*	He et al., 2008b
Resistance to Verticillium wilt[2]	(XinLuZao 1 × Hai7124) $F_{2:3}$ *	Wang et al., 2008a
Lint percentage and fiber quality traits[3]	(Yumian 1 × T586) F_2 and $F_{2:3}$	Zhang et al., 2005
Lint percentage, fiber quality and spiny bollworm[3]	(Yumian 1 × T586) F_2 and $F_{2:3}$	Wan et al., 2007
Fiber strength[3]	(CRI 8 × Pima90-53) F_2 *	Liu et al., 2009

[1]Published from Cotton Research Institute, Nanjing Agricultural University (CRI, NAU); [2]from Huazhong Agricultural University (HAU); [3]from other Universities. *Interspecific cross.

(ZMS 12 × J8891). This result was confirmed through further analyses (He et al., 2007, Qin et al., 2008b). Moreover, clustered QTLs for seed cotton and lint yield, specific yield components such as lint index, boll size, seed index, as well as fiber strength and micronaire, also were detected on D8 (Shen et al., 2006, 2007). These results indicate that genes controlling fiber development and yield can be linked, or that they are likely to be pleiotropic, resulting in negative relationships between fiber and yield components that cause complications for plant breeders.

Transcriptome analyses of fiber development

Cotton fiber is an excellent model for cellular development and elongation, which occurs in four overlapping stages: initiation, elongation, secondary cell wall (SCW) synthesis and maturation (Basra and Malik, 1984; Kim and Triplett, 2001). Cotton fiber initiation stage acts as a developmental switch to determine the number of fibers on each ovule, whereas the rate and duration of cell elongation/expansion determine fiber length, and the duration of SCW affects fiber strength and fineness (Smart et al., 1998; Ruan et al., 2001). Therefore, cotton fiber transcriptomics mostly focus on functional identification of crucial genes for improving fiber yield and quality. To date, many specifically or preferentially expressed genes have been identified in fiber (Table 3). Transcription factors play essential roles in cotton fiber initiation. Previous findings illustrate that complex networks (MYB–bHLH–WD40) control *Arabidopsis* trichome cell fate (Ramsay and Glover, 2005). The initiation of cotton fiber cells

was found to be developmentally similar to that of *Arabidopsis* trichomes (Guan et al., 2007). Therefore, identification of comparable transcription factors in cotton is very important for dissecting fiber initiation mechanisms. Functional analyses have demonstrated that *GaMYB2, GaRDL1, GaHOX1 and GhMYB109* play essential roles in the regulatory networks during cotton fiber initiation (Wang et al., 2004; Shangguan et al., 2008; Guan et al., 2008; Pu et al., 2008). Because cotton fiber is an excellent general model for cell elongation, the elongation phase is perhaps the best-studied period of fiber development (Kim and Triplett, 2001). Several cDNA libraries derived from ovules, fibers and other tissues from cultivated tetraploid cotton were constructed (Ji et al., 2003; Liu et al., 2006; Shi et al., 2006; Tu et al., 2007; Gou et al., 2007). Using a PCR-selected cDNA subtractive analysis and differential screening, 172 differentially expressed genes were identified between Xuzou142 fiber and fuzzless-lintless during fiber elongation (Ji et al., 2003); 292 preferentially expressed genes were identified between 10 and 20 days post anthesis (DPA) in fiber cells and non-fiber tissues (Liu et al., 2006), and 645 were identified between different stages of Pima3-79 ovules or fibers (Tu et al., 2007). A recently identified small regulatory RNAs, miRNAs are also differentially expressed during cotton fiber development (Zhang and Pan, 2009; Zhang et al., 2007). With implementation of large-scale EST sequencing, fiber elongation was studied using high-throughput DNA microarray platforms (Shi et al., 2006; Gou et al., 2007; Pang et al., 2010b). Shi et al. (2006) reported a 12K cDNA microarray platform (GEO accession: GPL2610) containing 11,962 uniESTs from 5-10 DPA Xuzhou142 fibers. They were the first to demonstrate that ethylene plays an essential role in promoting fiber cell elongation by activating fiber-specific genes, such as *SUS*, *EXP1*, *EXP2*, and *TUB1* that are important for cell wall biosynthesis, wall loosening and cytoskeleton rearrangement. Qin et al. (2007b) further demonstrated that very-long-chain fatty acids (*VLCFAs*) promote cotton fiber and *Arabidopsis* cell elongation by the activating ethylene biosynthesis gene *ACO*s. The *GhAPX1* gene has been shown to be involved in hydrogen peroxide (H_2O_2) homeostasis during cotton fiber development; H_2O_2 production is promoted by ethylene, and H_2O_2 induce ethylene production by a feedback regulatory mechanism, which together modulate cotton fiber development (Li et al., 2007; Qin et al., 2008a). Additionally, Luo et al. (2007) demonstrated that the steroid 5d-reductase (*GhDET2*) plays a crucial role in the initiation and elongation of cotton fiber cells, and that modulation of brassinosteroid (BR) biosynthesis factors can improve fiber quality or yield. It also was shown that gibberellin (GA) 20-oxidase (*GhGA20ox1-3*) promotes initiation and elongation of cotton fibers by regulating GA synthesis (Xiao et al., 2010). Based on this research on biosynthesis and signaling pathways, Mei et al. (2010) suggested a novel molecular mechanism of interactions among ethylene, BR, GA, H_2O_2 and *VLCFAs* during fiber cell elongation (Fig. 1A). In addition to biosynthesis of various phytohormones, fast polarized growth of a cotton fiber cell requires biosynthesis of plasma membrane and cell wall components, along with cell wall loosing and expansion. Functional analyses demonstrated that several genes related to the cytoskeleton (*GhTUB1, GhWBC1, GhPFN1, GhACT1, GhTUA9* and *GhTUB*s) and four genes related to cell wall biosynthesis or cell expansion (*GhRLK1, GhGlcAT1, GhPEL*, and *GhATPδ1*) play important roles in fiber elongation (Table 3). Gou et al. (2007) constructed a 5K cDNA array (GPL3641) covering 5,122 unique ESTs from a cDNA library of *G. hirsutum* L. cv. Xuzhou142 using -3 to 5 DPA ovules and 6-24 DPA fibers. They identified 633 differentially regulated genes during cell elongation and SCW synthesis, which indicated that auxin signaling, wall-loosening and lipid metabolism are highly active during fiber elongation, whereas cellulose biosynthesis is predominant in the SCW synthesis stage. Recently Pang et al. (2010b) reported a new cotton 32K cDNA microarray (GPL5476) containing 31,401 UniESTs. Large-scale cotton EST sequencing also provides a powerful platform for predicting microRNAs, which will increase our understanding mechanistic roles in regulating fiber development. Qiu et al. (2007) used bioinformatics approaches to identify microRNAs and their targets from the *G. hirsutum* ESTs database in NCBI, and Kwak et al. (2009) further enriched a set of microRNAs for fiber development. However, the role of small RNAs, especially microRNAs, in cotton fiber cell development is under-explored.

Transcriptome analyses of stress resistance

Both cotton growth and yield are severely inhibited by biotic and abiotic stresses. The complex stress response mechanism of cotton is being unraveled through the identification of stress response protein-encoding genes. Mao et al. (2007) made a major breakthrough, using RNA interference (RNAi) to improve stress resistance. They identified a cytochrome P450 gene (*CYP6AE14*) from cotton bollworm and silencing *CYP6AE14* by plant-mediated RNAi can impair larval tolerance of gossypol. It is reasonable to expect that, in the future, plant-mediated RNAi will be useful in producing transgenic cottons that are resistance to insects. Many differentially expressed ESTs involved in the defense response to *Verticillium* wilt have been identified (Zuo et al., 2005; Zhu et al., 2005; Gao et al., 2006). Ethylene-responsive element binding factors (ERFs/EREB) are plant-specific transcription factors, many of which could play dual roles in biotic and abiotic stresses. Meng et al. (2010) reported that the *EREB1* gene might play an important role in *V. dahliae* stress signal transduction pathways by activating pathogenesis-related genes. Yang et al. (2010) reported cytochrome P450 reductase (*GhCPR2*) transcription was induced dramatically by *V. dahliae* elicitor in suspension culture of cotton cells, and was more related to defense reactions. Further studies should be performed to clarify the role of the *GhCPR2* response to *V. dahliae*. These genes could facilitate breeding of *V. dahliae*-resistant cotton varieties in future. Molecular studies of abiotic stresses will be helpful for improving tolerant cotton varieties. Ethylene-responsive factor genes (*GhERF1-4,6*) were isolated from Upland cotton responses to multiple abiotic stresses (Jin and Liu, 2008; Qiao et al., 2008; Jin et al., 2010). Two other gene family were identified that respond to abiotic stresses, including nineteen novel cotton fasciclin-like arabinogalactan protein genes related to salt stress (Huang et al., 2008b) and six novel *NAC* genes that respond to drought, cold and ABA stress (Meng et al., 2009). *GhDBP2* (DRE-binding protein) was involved in responses to environmental stresses as well as ABA treatment (Huang et al., 2008a). Recently, Xue et al. (2009) reported that accumulation of mRNA for a 64-amino acid type 3 metallothionein protein (*GhMT3a*) up-regulated by ABA, ethylene and reactive oxygen species (ROS) in cotton seedlings, indicating that *GhMT3a* could function as an effective ROS scavenger and that its expression could be regulated by abiotic stresses through ROS signaling. These results have helped to deepen our understanding of the molecular mechanisms of cotton resistance stress, and they identify candidate genes for improving cotton resistance and/or increased tolerance to stress *via* genetic engineering strategies.

Table 3. A survey of cotton fiber specifically or preferentially expressed genes.

Stages	Gene name	Putative functions during cotton fiber development	References
Fiber initiation	*GaMYB2(FIF1)*	Predominantly expressed early in the development of cotton fibers; and rescued the trichome formation of *Arabidopsis gl1* mutant.	Wang et al., 2004; Shangguan et al., 2008
	GaRDL1	Contains a homeodomain binding L1 box involved in activating the *RDL1-P3* promoter in *Arabidopsis* trichomes, and *RDL1* was expressed mainly in developing fiber cells.	Wang et al.,2004
	GaHOX1	*GaHOX1* is predominantly expressed in cotton fiber cells at early developmental stages, and is a functional homolog of GL2 in plant trichome development.	Guan et al., 2008
Initiation& elongation	*GhMYB109*	*GhMYB109* is specifically expressed in cotton initial and elongating fibers and revealed a largely conserved mechanism of the R2R3 MYB transcription factor in cell fate determination in plants.	Suoet al.,2003;Pu et al., 2008
	GhDET2	*GhDET2* plays a crucial role in the initiation and elongation of cotton fiber cells.	Luo et al., 2007
	GhFLA1	The *FLAs* are essential for the initiation and elongation of cotton fiber development.	Liu et al., 2008; Li et al., 2010
	GhAGP2,3,4		
	GhGA20ox1-3	*GhGA20ox1* is expressed preferentially in elongating fiber, while *GhGA20ox2-3* transcripts accumulate mainly in ovules; they promote initiation and elongation of cotton fiber by regulating gibberellins synthesis.	Xiao et al., 2010
Fiber elongation	*GhKCR1-2*	Encoding 3-ketoacyl-CoA reductases, and preferentially expressed during cotton fiber elongation, *GhKCR1* and *GhKCR2* play an important role in very long chain fatty acids biosynthesis.	Qin et al., 2005
	GhKCS13/CER6	Encoding 3-ketoacyl-CoA synthase, involved in *VLCFAs* (very-long-chain fatty acids) biosynthesis; *VLCFAs* promote cotton fiber and *Arabidopsis* cell elongation by activating ethylene biosynthesis.	Qin et al., 2007a, 2007b
	GhECR1-2	Encoding trans-2-enoyl-CoA reductase (*ECR*), and has up-regulated expression during fiber elongation, involved in fatty acid elongation during cotton fiber development.	Song et al., 2009a
	GhAPX1	*GhAPX1* has up-regulated expression in response to an increase in cellular H_2O_2 and ethylene, and encodes a functional enzyme involved in hydrogen peroxide homeostasis during fiber development.	Li et al., 2007; Qin et al., 2008a
	GhWBC1	Encodes an ATP-binding cassette transporter of the WBC subfamily with highly expression in developing fiber cells, and over-expressed *GhWBC1* interferes with *Arabidopsis* seed and silique development.	Zhu et al., 2003
	GhACT1	Encodes an actin and is involved in fiber elongation, but not in fiber initiation.	Li et al., 2005a
	GhTUB1	Encoding β-tubulin, and preferentially accumulating at high levels in fiber, may play a distinct and required role in fiber development.	Li et al., 2002
	GhTUBs	Nine *GhTUBs* were highly expressed in elongating fiber cells as compared with fuzzless-lintless mutant ovules, and were induced by gibberellin, ethylene, brassinosteroids, and lignoceric acid.	He et al., 2008a
	GhTUA9	*GhTUA9* gene is specifically expressed in fiber and involved in cell elongation.	Li et al., 2007
	GhPFN1	May be involved in the rapid elongation of cotton fibers by promoting actin polymerization.	Wang et al., 2005
	Gh14-3-3L	*Gh14-3-3L* is predominantly expressed during early fiber development, and reaches peak of expression in 10 DPA fiber cells involved in regulating fiber elongation.	Shi et al., 2007; Zhang et al., 2010
	GhBG	*GhBG* is highly abundant in 5-17 DPA fiber and can lead to a significant increase in cell length and width when transformed into yeast.	Ma et al., 2006
	GhPEL	Encoding a pectate lyase, may block cell wall loosening by depolymerization of de-esterified pectin during fiber elongation.	Wang et al., 2010a
	GhEF1As	Translation elongation factor 1A-1, 2, 4, 5 and 9 active at the early fiber elongation.	Xu et al., 2007
	GhATPδ1	*GhATPδ1* (ATP synthase δ1 subunit) is important for activity of mitochondrial ATP synthase, probably relates to fiber elongation.	Pang et al., 2010a
	GhGS	*GhGS* is differentially expressed between 7235 and TM-1 at 8 DPA fibers, significantly correlated with fiber strength QTL on D7.	He et al.,2008c
Elongation & SCW	*GhGlcAT1*	*GhGlcAT1* may be involved in non-cellulose polysaccharides biosynthesis of the cotton cell wall.	Wu et al., 2005, 2007
	GhRLK1	*GhRLK1* is expressed in fast-elongation and the transition stage of elongation and SCW, and involve in the induction and maintenance of active fiber secondary wall formation.	Li et al., 2005b

Proteome analyses of fiber development

Gene expression at the mRNA level does not reveal exact functions of genes in cells; therefore, direct research on protein expression patterns and functional models has become an inevitable trend in life sciences. The term proteome was coined to describe the set of proteins encoded by a given genome (Wilkins et al., 1996). Protein profiling is one of the important recent developments in proteomics; it offers multiple advantages and complements other functional genomics approaches such as transcript profiling. After an extraction protocol for 2-D electrophoresis (2-DE) was optimized (Yao et al., 2006), a proteomic analysis of cotton fibers during cell elongation was conducted (Yang et al., 2008). It identified differentially expressed proteins from mass spectrometry, which match 66 unique protein species involved in different cellular and metabolic processes, with obvious functional tendencies toward energy/carbohydrate metabolism, protein turnover, cytoskeleton dynamics, cellular responses and redox homeostasis. This provides a global view of the development-dependent protein changes in cotton fibers, and offers a framework for further functional research that targets proteins associated with fiber development. Using a comparative proteomics approach, Pang et al. (2010) identified 104 proteins from 10 DPA cotton ovules, with 93 preferentially accumulating in the wild-type and 11 accumulating in the

fuzzless-lintless mutant, and identified nucleotide sugar metabolism as the most significantly up-regulated biochemical process during fiber elongation. Seven protein spots potentially involved in pectic cell wall polysaccharide biosynthesis specifically accumulated in wild-type samples at both protein and transcript levels. Comparative proteomics indicate that biosynthesis of pectic precursors is important for cotton fiber and *Arabidopsis* root hair elongation (Fig. 1B). Zhao et al. (2010) identified 81 differentially expressed proteins from Ligon lintless (Li_1) fibers assigned to different functional categories through 2-DE combined with local EST database-assisted MS/MS analysis; 54 of these proteins were down-regulated and 27 up-regulated. Of these, over half of the down-regulated proteins are mainly involved in protein folding and stabilization, nucleocytoplasmic transport, signal transduction, and vesicular-mediated transport and a number of cytoskeleton-related proteins showed a remarkable decrease in protein abundance in the Li_1 fibers. Accordingly, the architecture of the actin cytoskeleton was severely deformed and microtubule organization was moderately altered, accompanied by dramatic disruption of vesicle trafficking. By contrast, the expression of several proteins involved in unfolded protein response was activated in Li_1 fibers, which indicated that the deficiency of fiber cell elongation was related to endoplasmic reticulum (ER) stress. Collectively, these findings significantly enhance our understanding of mechanisms associated with cotton fiber elongation.

Future prospects

In recent decades, new tools of transcriptome analysis in China have been applied to cotton including cDNA-amplified fragment length polymorphism (cDNA-AFLP) (Pan et al., 2007; Liu et al., 2009; Zhu et al., 2009), microarrays (Shi et al., 2006; Gou et al., 2007), and next-generation sequencing (NGS). cDNA-AFLP is a PCR-based transcript profiling technology that does not require any prior knowledge of gene sequences, and combines the advantage of high specificity with the capability of detection of rare transcript tags; therefore, its sensitivity is higher than that of hybridization-based techniques. Conversely, the strength of microarrays lies in their massive parallel nature, allowing the simultaneous analysis of up to tens of thousands of genes. Recently, superior, higher capability genome-wide NGS platforms (e.g. Roche (454), Illumina and SOLiD) have been developed, and presumably will accelerate advances in genomics and transcriptomics dramatically (Shendure et al., 2008). NGS should become inexpensive, routine and widespread for studies of the genomes and transcriptomes in the near future. In reviewing the status of cotton omics, it is clear that Chinese scientists have made significant progress in fields of constructing genetic maps, genes or QTL mapping, transcriptome analysis of fiber, despite the fact that our omics research was launched later than in developed countries. Nevertheless, many efforts are needed to further develop omics resources and approaches in order to fully and effectively use them in cotton genetic improvement and biological research. In particular, the following areas of cotton omics research should be emphasized.

Transcriptomics, proteomics and metabolomics

The attractiveness of cotton as a model of single cell development has been acknowledged, and efforts are underway worldwide to elucidate genetic features that are key to generating superior fiber species. In the post-genomic era, various studies have focused on connecting gene function and gene expression with resulting phenotype through complex networks of DNA → RNA → protein → metabolite → phenotype. A large number of studies have demonstrated that cotton fiber development involves complex molecular mechanisms, and cotton fiber cell activities require complex patterns of gene transcription, protein expression as well as related metabolic pathways. Therefore cotton proteomics and metabolomics are important directions for post-genomics cotton research aimed at understanding molecular mechanisms of cotton fiber development, because they bridge roles between gene expression and phenotype. Using the NGS technology for RNA profiling, we can discover more and novel tags that are differentially expressed. Thus, in order to dissect cotton fiber developmental mechanisms more deeply, continued efforts should be made in transcriptomics, proteomics and metabolomics.

Integrating omics

Future directions also will include the integration of different omics in cotton fiber development. The trend in biological investigations is shifting from individual omics toward integrated omics and system biology. Integration of molecular profiling technologies into plant developmental biology has just begun, and many exciting developments can be anticipated in the near future (Hennig, 2007). Gou et al. (2007) have developed a preliminary transcriptome integrated with metabolome in cotton fiber development studies, and demonstrated that signaling and metabolic pathways are coordinated to promote cell elongation in the early stage and to support cellulose synthesis in later stages. Therefore, with high-throughput data acquisition by genomic projects, it is possible and necessary to better integrate multi-omics technologies and systems approaches that will generate many intriguing insights into cotton fiber development.

Quantitative gemomics

Quantitative genetics in the age of omics will expand in cotton. Genetical genomics, which combines genetics with large scale expression profiling to provide expression QTLs (eQTLs), has been applied in *Arabidopsis*, maize, and barley (Druka et al., 2010). Similar approaches can be followed with data derived from other "omics" technologies such as proteomics (pQTLs) and metabolomics (mQTLs) (Keurentjes et al., 2008; Joosen et al., 2009). Genetic regulatory networks have shown the usefulness of combining quantitative genetics and large-scale omics analyses (Keurentjes et al., 2007, 2008). Using these approaches, we will be able to integrate genetic, transcriptomic, proteomic and metabolomic data (eQTL, pQTL and mQTL) to understand molecular mechanisms and constructing regulatory networks that underlie complex cotton fiber qualities.

Genomics-assisted breeding

With various current and developing "omics" technologies, marker-assisted breeding and selection will gradually evolve into genomics-assisted breeding for crop improvement. Eventually, knowledge of the relative values of alleles at all segregating loci in a population could allow the breeder to design a genotype *in silico* and to practice whole genome selection (Varshney et al., 2005). The Upland cotton genome sequencing project can be enhanced through the use of NGS technology, which will enable us to discovery a large number of single-nucleotide polymorphisms (SNPs) within whole genome sequences or large genomic fragments in BACs that can be applied to genome-wide association (GWS) study as in *Arabidopsis* (Atwell et al., 2010). This can help to identify

genetic loci or genes associated with traits of agricultural importance. Genomics-assisted breeding will be an effective approach to overcoming the bottlenecks of conventional breeding practices, through the integration of germplasm resources, genetic and genomic resources, and multiple omics tools and strategies. Ultimately this will lead to improvements in cotton fiber yield, quality and pest resistance.

Acknowledgment

We thank Dr. RJ Kohel for his critical editing the manuscript. We also thank all faculties and graduate students in Cotton Research Institute, Nanjing Agricultural University and the other members of the cotton genomics and breeding community in China for their contributions, and apologize for not citing many enlightening papers owing to space limitations. This study was supported by grants from the National Nature Science Foundation of China (Grant No.30730067) and Nature Science Foundation in Jiangsu Province (Grant No. BK2008036).

References

Atwell S, Huang YS, Vilhjalmsson BJ, Willems G, Horton M, Li Y, Meng D, Platt A, Tarone AM, Hu TT, Jiang R, Muliyati NW, Zhang X, Amer MA, Baxter I, Brachi B, Chory J, Dean C, Debieu M, de Meaux J, Ecker JR, Faure N, Kniskern JM, Jones JD, Michael T, Nemri A, Roux F, Salt DE, Tang C, Todesco M, Traw MB, Weigel D, Marjoram P, Borevitz JO, Bergelson J, Nordborg M (2010) Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. Nature 465: 627-631

Basra AS, Malik CP (1984) Development of the cotton fiber. International Review of Cytology 89: 65-113

Chen DY, Ding YZ, Guo WZ, Zhang TZ (2009a) Molecular mapping of genic male-sterile genes *ms15*, *ms5* and *ms6* in tetraploid cotton. Plant Breeding 128 (2):193-198

Chen H, Qian N, Guo WZ, Song QP, Li BC, Deng FJ, Dong CG, Zhang TZ (2009b) Using three overlapped RILs to dissect genetically clustered QTL for fiber strength on Chro.D8 in Upland cotton. Theor Appl Genet 119: 605-612

Dong CG, Ding YZ, Guo WZ, Zhang TZ (2007) Fine mapping of the dominant glandless Gene Gl^e_2 in Sea-island cotton (*Gossypium barbadense* L.) Chinese Sci Bull 52: 3105-3109

Druka A, Potokina E, Luo ZW, Jiang N, Chen XW, Kearsey M, Waugh R (2010) Expression quantitative trait loci analysis in plants. Plant Biotechnology Journal 8: 10-27

Gao YL, Guo WZ, Wang L, Zhang TZ (2006) Isolation and characterization of resistance and defense gene analogs in cotton (*Gossypium barbadense* L.). Sci China C Life Sci 49: 530-542

Gou JY, Wang LJ, Chen SP, Hu WL, Chen XY (2007) Gene expression and metabolite profiles of cotton fiber during cell elongation and secondary cell wall synthesis. Cell Research 17: 422-434

Guan XY, Li QJ, Shan CM, Wang S, Mao YB, Wang LJ, Chen XY (2008) The HD-Zip IV gene *GaHOX1* from cotton is a functional homologue of the *Arabidopsis GLABRA2*. Physiol Plant 134: 174-182

Guan XY, Yu Na, Shangguan XX, Wang S, Lu S, Wang LJ, Chen XY (2007) *Arabidopsis* trichome research sheds light on cotton fiber development mechanisms. Chinese Sci Bull 52: 1734-1741

Guo WZ, Cai CP, Wang CB, Han ZG, Song XL, Wang K, Niu XW, Wang C, Lu KY, Shi B, Zhang TZ (2007) A microsatellite-based, gene-rich linkage map reveals genome structure, function and evolution in *Gossypium*. Genetics 176: 527-541

Guo WZ, Cai CP, Wang CB, Zhao L, Wang L, Zhang TZ (2008) A preliminary analysis of genome structure and composition in *Gossypium hirsutum*. BMC Genomics 9: 314

Guo WZ, Zhang TZ, Pan JJ, Kohel RJ (1997) Screening the RAPD-PCR marker of cotton fertility restoration genes in cytoplasmic male sterile lines. Chinese Sci Bull 42 (24):2645-2647 (in Chinese)

Guo WZ, Zhang TZ, Shen XL, Yu JZ, Kohel RJ (2003) Development of SCAR marker linked to a major QTL for high fiber strength and its usage in molecular-marker assisted selection in upland cotton. Crop Science 43(6):2252–2256

Han ZG, Guo WZ, Song XL, Zhang TZ (2004) Genetic mapping of EST-derived microsatellites from the diploid *Gossypium arboreum* in allotetraploid cotton. Mol Genet Genomics 272: 308-327

Han ZG, Wang CB, Song XL, Guo WZ, Gou JY, Li CH, Chen XY, Zhang TZ (2006) Characteristics, development and mapping of *Gossypium hirsutum* derived EST-SSRs in allotetraploid cotton. Theor Appl Genet 112: 430-439

He DH, Lin ZX, Zhang XL, Nie YC, Guo XP, Zhang YX (2007) QTL mapping for economic traits based on a dense genetic map of cotton with PCR-based markers using the interspecific cross of *Gossypium hirsutum* × *G. barbadense*. Euphytica 153:181-197

He DH, Lin ZX, Zhang XL, Zhang YX, Li W, Nie YC, Guo XP (2008b) Dissection of genetic variance of fiber quality in advanced generations from an interspecific cross of *Gossypium hirsutum* and *G. barbadense*. Plant Breeding 127: 286-294

He DH, Lin ZX, Zhang XL，Nie YC, Guo XP, Feng CD, Stewart JM (2005) Mapping QTLs of traits contributing to yield and analysis of genetic effects in tetraploid cotton. Euphytica 144(1-2):141-149

He XC, Qin YM, Xu Y, Hu CY, Zhu YX (2008a) Molecular cloning, expression profiling, and yeast complementation of 19 beta-tubulin cDNAs from developing cotton ovules. J Exp Bot 59: 2687-2695

He YJ, Guo WZ, Shen XL, Zhang TZ (2008c) Molecular cloning and characterization of a cytosolic glutamine synthetase gene, a fiber strength-associated gene in cotton. Planta 228: 473-483

Hennig L (2007) Patterns of beauty–omics meets plant development. Trends Plant Sci 12(7): 287-293

Hu Y, Guo WZ, Zhang TZ (2009) Construction of a bacterial artificial chromosome library of TM-1, a standard line for genetics and genomics in Upland cotton. J Integr Plant Biol 51: 107-112

Huang B, Jin L, Liu JY (2008a) Identification and characterization of the novel gene *GhDBP2* encoding a DRE-binding protein from cotton (*Gossypium hirsutum*). J Plant Physiol 165: 214-223

Huang GQ, Xu WL, Gong SY, Li B, Wang XL, Xu D, Li XB (2008b) Characterization of 19 novel cotton *FLA* genes and their expression profiling in fiber development and in response to phytohormones and salt stress. Physiol Plant 134: 348-359

Ji SJ, Lu YC, Feng JX, Wei G, Li J, Shi YH, Fu Q, Liu D, Luo JC, Zhu YX (2003) Isolation and analyses of genes preferentially expressed during early cotton fiber development by subtractive PCR and cDNA array. Nucleic Acids Res 31: 2534-2543

Jiang F, Zhao J, Zhou L, Guo WZ, Zhang TZ (2009) Molecular mapping of Verticillium wilt resistance QTL clustered on chromosomes D7 and D9 in upland cotton. Sci China C Life Sci 52: 872-884

Jin LG, Li H, Liu JY (2010) Molecular characterization of three ethylene responsive element binding factor genes from cotton. J Integr Plant Biol 52: 485-495

Jin LG, Liu JY (2008) Molecular cloning, expression profile and promoter analysis of a novel ethylene responsive transcription factor gene *GhERF4* from cotton (*Gossypium hirsutum*). Plant Physiol Biochem 46: 46-53

Joosen RV, Ligterink W, Hilhorst HW, Keurentjes JJ (2009) Advances in Genetical Genomics of Plants. Current Genomics 10: 540-549

Keurentjes JJ, Fu J, Terpstra IR, Garcia JM, van den Ackerveken G, Snoek LB, Peeters AJ, Vreugdenhil D, Koornneef M, Jansen RC (2007) Regulatory network construction in Arabidopsis by using genome-wide gene expression quantitative trait loci. Proc Natl Acad Sci USA 104:1708-1713

Keurentjes JJ, Koornneef M, Vreugdenhil D (2008) Quantitative genetics in the age of omics. Current Opinion in Plant Biology 11:123–128

Kim HJ, Triplett BA (2001) Cotton fiber growth in planta and in vitro. Models for plant cell elongation and cell wall biogenesis. Plant Physiol 127(4):1361-6.

Kwak PB, Wang QQ, Chen XS, Qiu CX, Yang ZM (2009) Enrichment of a set of microRNAs during the cotton fiber development. BMC Genomics 10:457

Li HB, Qin YM, Pang Y, Song WQ, Mei WQ, Zhu YX (2007) A cotton ascorbate peroxidase is involved in hydrogen peroxide homeostasis during fibre cell development. New Phytol 175: 462-471

Li XB, Cai L, Cheng NH, Liu JW (2002) Molecular characterization of the cotton *GhTUB1* gene that is preferentially expressed in fiber. Plant Physiol 130: 666–674

Li XB, Fan XP, Wang XL, Cai L, Yang WC (2005a) The cotton *ACTIN1* gene is functionally expressed in fibers and participates in fiber elongation. Plant Cell 17: 859-875

Li YJ, Liu DQ, Tu LL, Zhang XL, Wang L, Zhu LF, Tan JF, Deng FL (2010) Suppression of *GhAGP4* gene expression repressed the initiation and elongation of cotton fiber. Plant Cell Rep 29(2):193- 202

Li YL, Sun J, Xia GX (2005b) Cloning and characterization of a gene for an LRR receptor-like protein kinase associated with cotton fiber development. Mol Genet Genomics 273: 217-224

Lin ZX, He DH, Zhang XL, Nie YC, Guo XP, Feng CD, Stewart M (2005) Linkage map construction and mapping QTL for cotton fiber quality using SRAP, SSR and RAPD. Plant Breeding 124:180-187.

Lin ZX, Zhang XL, Nie YC, He DH, Wu MQ (2003) Construction of a genetic linkage map for cotton based on SRAP. Chinese Sci Bul 48(19):2063–2067

Lin ZX, Zhang YX, Zhang XL, Guo XP (2009) A high-density integrative linkage map for *Gossypium hirsutum* L. Euphytica 166(1): 35-45

Liu DQ, Tu LL, Li YJ, Wang L, Zhu LF, Zhang XL (2008) Genes encoding fasciclin-like arabinogalactan proteins are specifically expressed during cotton fiber development. Plant Mol Bio Rep 26(2): 98-113

Liu DQ, Zhang XL, Tu LL, Zhu LF, Guo XP (2006) Isolation by suppression-subtractive hybridization of genes preferentially expressed during early and later fiber development stages in cotton. Molecular Biology, 40(5): 741-749

Liu HW, Wang XF, Pan YX, Shi RF, Zhang GY, Ma ZY (2009) Mining cotton fiber strength candidate genes based on transcriptome mapping. Chinese Sci Bul 54: 4651-4657

Liu LW, Guo WZ, Zhu XF, Zhang TZ (2003) Inheritance and fine mapping of fertility restoration for cytoplasmic male sterility in *Gossypium hirsutum* L. Theor Appl Genet 106(3):461–469

Luo M, Xiao YH, Li XB, Lu XF, Deng W, Li DM, Hou L, Hu MY, Li Y, Pei Y (2007) *GhDET2*, a steroid 5alpha-reductase, plays an important role in cotton fiber cell initiation and elongation. Plant J 51: 419-430

Ma GJ, Zhang TZ, Guo WZ (2006) Cloning and characterization of cotton *GhBG* gene encoding beta-glucosidase. DNA Seq 17: 355-362

Ma XX, Ding YZ, Zhou BL, Guo WZ, Lv YH, Zhu XF, Zhang TZ (2008b) QTL mapping in A-genome diploid Asiatic cotton and their congruence analysis with AD-genome tetraploid cotton in genus *Gossypium*. J Genet Genomics 35: 751-762

Ma XX, Zhou BL, Lu YH, Guo WZ, Zhang TZ (2008a) Simple sequence repeat genetic linkage maps of A-genome diploid cotton (*Gossypium arboreum*). J Integr Plant Biol 50: 491-502

Mao YB, Cai WJ, Wang JW, Hong GJ, Tao XY, Wang LJ, Huang YP, Chen XY (2007) Silencing a cotton bollworm P450 monooxygenase gene by plant-mediated RNAi impairs larval tolerance of gossypol. Nat Biotechnol 25: 1307-1313

Mei WQ, Qin YM, Zhu YX (2010) Cross-talk among ethylene, very long-chain fatty acids, reactive oxygen species, brassinosteroid and gibberellin mediates cotton fiber elongation. Chinese Bulletin of Life Sciences 22(1):7-14 (in Chinese)

Meng CM, Cai CP, Zhang TZ, Guo WZ (2009) Characterization of six novel *NAC* genes and their responses to abiotic stresses in *Gossypium hirsutum* L. Plant Science 176 (3): 352-359

Meng XP, Li FG, Liu CL, Zhang CJ, Wu ZX, Chen YJ (2010) Isolation and characterization of an ERF transcription factor gene from cotton (*Gossypium barbadense* L.). Plant Mol Biol Rep 28:176–183

Oksman-Caldentey KM, Saito K (2005) Integrating genomics and metabolomics for engineering plant metabolic pathways. Curr Opin Biotechnol 16: 174-179

Pan YX, Ma J, Zhang GY, Han GY, Wang XF, Ma ZY (2007) cDNA-AFLP profiling for the fiber development stage of secondary cell wall synthesis and transcriptome mapping in cotton. Chinese Sci Bul 52(17):2358-2364

Pang CY, Wang H, Pang Y, Xu C, Jiao Y, Qin YM, Western TL, Yu SX, Zhu YX (2010b) Comparative proteomics indicate that biosynthesis of pectic precursors is important for cotton fiber and *Arabidopsis* root hair elongation. Mol Cell Proteomics 9:2019-2033

Pang CY, Wang H, Song WQ, Zhu YX (2010a) The cotton ATP synthaseδ1 subunit is required to maintain a higher ATP⁄ADP ratio that facilitates rapid fibre cell elongation. Plant Biology 12(6):903-9

Pu L, Li Q, Fan XP, Yang WC, Xue YB (2008) The R2R3 MYB transcription factor *GhMYB109* is required for cotton fiber development. Genetics 180: 811-820

Qian N, Zhang XW, Guo WZ, Zhang TZ (2009) Fine mapping of open-bud duplicate genes in homoelogous chromosomes of tetraploid cotton. Euphytica 165: 325-331

Qiao ZX, Huang B, Liu JY (2008) Molecular cloning and functional analysis of an ERF gene from cotton (*Gossypium hirsutum*). Biochim Biophys Acta 1779: 122-127

Qin HD, Guo WZ, Zhang YM, Zhang TZ (2008b) QTL mapping of yield and fiber traits based on a four-way cross population in *Gossypium hirsutum* L. Theor Appl Genet 117: 883-894

Qin YM, Hu CY, Pang Y, Kastaniotis AJ, Hiltunen JK, Zhu YX (2007b) Saturated very-long-chain fatty acids promote cotton fiber and *Arabidopsis* cell elongation by activating ethylene biosynthesis. Plant Cell 19: 3692-3704

Qin YM, Hu CY, Zhu YX (2008a) The ascorbate peroxidase regulated by H_2O_2 and ethylene is involved in cotton fiber cell elongation by modulating ROS homeostasis. Plant Signal Behav 3: 194-196

Qin YM, Pujol FM, Hu CY, Feng JX, Kastaniotis AJ, Hiltunen JK, Zhu YX (2007a) Genetic and biochemical studies in yeast reveal that the cotton fibre-specific *GhCER6* gene functions in fatty acid elongation. J Exp Bot 58: 473-481

Qin YM, Pujol FM, Shi YH, Feng JX, Liu YM, Kastaniotis AJ, Hiltunen JK, Zhu YX (2005) Cloning and functional characterization of two cDNAs encoding NADPH-dependent 3-ketoacyl-CoA reductased from developing cotton fibers. Cell research 15: 465-473

Qin YM, Zhu YX (2007c) A brief summary of major advances in cotton functional genomics and molecular breeding studies in China. Chin Sci Bull 52(23):3174-3178

Qiu CX, Xie FL, Zhu YY, Guo K, Huang SQ, Nie L, Yang ZM (2007) Computational identification of microRNAs and their targets in *Gossypium hirsutum* expressed sequence tags. Gene 395:49–61

Ramsay NA, Glover BJ (2005) MYB-bHLH-WD40 protein complex and the evolution of cellular diversity. Trends Plant Sci 10: 63-70

Ruan YL, Llewellyn DJ, Furbank RT (2001) The control of single-celled cotton fiber elongation by developmentally reversible gating of plasmodesmata and coordinated expression of sucrose and K^+ transporters and expansin. Plant Cell 13(1):47-60.

Shangguan XX, Xu B, Yu ZX, Wang LJ, Chen XY (2008) Promoter of a cotton fibre MYB gene functional in trichomes of Arabidopsis and glandular trichomes of tobacco. J Exp Bot 59: 3533-3542

Shen XL, Guo WZ, Lu QX, Zhu XF, Yuan YL, Zhang TZ (2007) Genetic mapping of quantitative trait loci for fiber quality and yield trait by RIL approach in Upland cotton. Euphytica 155: 371-380.

Shen XL, Guo WZ, Zhu XF, Yuan YL, Yu JZ, Kohel RJ, Zhang TZ (2005) Molecular mapping of QTLs for qualities in three diverse lines in Upland cotton using SSR markers. Mol Breeding 15:169-181.

Shen XL, Zhang TZ, Guo WZ, Zhu XF, Zhang XY (2006) Mapping fiber and yield QTLs with main, epistatic, and QTL × environment interaction effects in recombinant inbred lines of Upland cotton. Crop Science 46(1): 61–66

Shendure J, Ji H (2008) Next-generation DNA sequencing. Nature Biotechnol 26(10):1135-1145

Shi HY, Wang XL, Li DD, Tang WK, Wang H, Xu WL, Li XB (2007) Molecular characterization of cotton *14-3-3L* gene preferentially expressed during fiber elongation. J Genet Genomics 34: 151-159

Shi YH, Zhu SW, Mao XZ, Feng JX, Qin YM, Zhang L, Cheng J, Wei LP, Wang ZY, Zhu YX (2006) Transcriptome profiling, molecular biological, and physiological studies reveal a major role for ethylene in cotton fiber cell elongation. Plant Cell 18: 651-664

Smart LB, Vojdani, F, Maeshima M, Wilkins TA (1998) Genes involved in Osmoregulation during turgor-driven cell expansion of developing cotton fibers are diferentially regulated. Plant Physiol 116: 1539–1549.

Song L, Guo WZ, Qin HD, Ding YZ, Zhang TZ (2010) Genetic analysis and molecular validation of chromosome assignment for fuzzless genes N_1 and n_2 in cotton. Journal of Nanjing Agricultural University 33(1):21-26 (In Chinese)

Song L, Guo WZ, Zhang TZ (2009b) Interaction of novel Dobzhansky-Muller type genes for the induction of hybrid lethality between *Gossypium hirsutum* and *G. barbadense* cv. Coastland R4-4. Theor Appl Genet 119: 33-41

Song WQ, Qin YM, Saito M, Shirai T, Pujol FM, Kastaniotis AJ, Hiltunen JK, Zhu YX (2009a) Characterization of two cotton cDNAs encoding trans-2-enoyl-CoA reductase reveals a putative novel NADPH-binding motif. J Exp Bot 60: 1839-1848

Song XL, Guo WZ, Han ZG, Zhang TZ (2005a) Quantitative trait loci mapping of leaf morphological traits and chlorophyll content in cultivated tetraploid cotton. J Integr Plant Biol 47(11): 1382−1390

Song XL, Wang K, Guo WZ, Zhang J, Zhang TZ (2005b) A comparison of genetic maps constructed from haploid and BC_1 mapping populations from the same crossing between *Gossypium hirsutum* L. and *Gossypium barbadense* L. Genome 48: 378-390

Song XL, Zhang TZ (2007) Identification of quantitative trait loci controlling seed physical and nutrient traits in cotton. Seed Science Research 17:243–251

Song XL, Zhang TZ (2009c) Quantitative trait loci controlling plant architectural traits in cotton. Plant Science 177 (4): 317-323

Suo JF, Liang XE, Pu L, Zhang YS, Xue YB (2003) Identification of *GhMYB109* encoding a R2R3 MYB transcription factor that expressed specifically in fiber initials and elongating fibers of cotton (*Gossypium hirsutum* L.). Biochim Biophys Acta 1630: 25-34

Tu LL, Zhang XL, Liang SG, Liu DQ, Zhu LF, Zeng FC, Nie YC, Guo XP, Deng FL, Tan JF, Xu L (2007) Genes expression analyses of Sea-island cotton (*Gossypium barbadense* L.) during fiber development. Plant Cell Rep 26: 1309–1320.

Varshney RK, Graner A, Sorrells ME (2005) Genomics-assisted breeding for crop improvement. Trends Plant Sci 10: 621-630

Wan Q, Zhang ZS, Hu MC, Chen L, Liu DJ, Chen X, Wang W, Zheng J (2007) T1 locus in cotton is the candidate gene affecting lint percentage, fiber quality and spiny bollworm (*Earias* spp.) resistance. Euphytica 158:241–247

Wang BH, Guo WZ, Zhu XF, Wu YT, Huang NT, Zhang TZ (2006c) QTL mapping for fiber quality in an elite hybrid derived-RIL population of upland cotton. Euphytica 152(3):367-378

Wang BH, Guo WZ, Zhu XF, Wu YT, Huang NT, Zhang TZ (2007a) QTL mapping of yield and yield components for elite hybrid derived-RILs in upland cotton. J Genet Genomics 34: 35-45

Wang BH, Wu YT, Guo WZ, Zhu XF, Huang NT, Zhang TZ (2007d) QTL analysis and epistasis effects dissection of fiber qualities in an elite cotton hybrid grown in second generation. Crop Science 47:1384–1392

Wang BH, Wu YT, Huang NT, Zhu XF, Guo WZ, Zhang TZ (2006a) QTL mapping for plant architecture traits in upland cotton using RILs and SSR markers. Acta Genetica Sinica 33: 161-170

Wang HH, Guo Y, Lv FN, Zhu HY, Wu SJ, Jiang YJ, Li FF, Zhou BL, Guo WZ, Zhang TZ (2010a) The essential role of *GhPEL* gene, encoding a pectate lyase, in cell wall loosening by depolymerization of the de-esterified pectin during fiber elongation in cotton. Plant Mol Biol 72: 397-406

Wang HM, Lin ZX, Zhang XL, Chen W, He DH, Guo XP, Nie YC, Li YH (2008a) Mapping and quantitative trait loci analysis of Verticillium wilt resistance genes in cotton. J Integr Plant Biol 50:174-182

Wang HY, Yu Y, Chen ZL, Xia GX (2005) Functional characterization of *Gossypium hirsutum* profilin 1 gene (*GhPFN1*) in tobacco suspension cells. Characterization of in vivo functions of a cotton profilin gene. Planta 222: 594-603

Wang K, Guan B, Guo WZ, Zhou BL, Hu Y, Zhu YC, Zhang TZ (2008b) Completely distinguishing individual A-genome chromosomes and their karyotyping analysis by multiple bacterial artificial chromosome - fluorescence in situ hybridization. Genetics 178: 1117-1122

Wang K, Guo WZ, Yang ZJ, Hu Y, Zhang WP, Zhou BL, Stelly DM, Chen ZJ, Zhang TZ (2010b) Structure and size variations between 12A and 12D homoeologous chromosomes based on high-resolution cytogenetic map in allotetraploid cotton. Chromosoma 119: 255-266

Wang K, Guo WZ, Zhang TZ (2007b) Detection and mapping of homologous and homoeologous segments in homoeologous groups of allotetraploid cotton by BAC-FISH. BMC Genomics 8:178

Wang K, Guo WZ, Zhang TZ (2007c) Development of one set of chromosome-specific microsatellite-containing BACs and their physical mapping in *Gossypium hirsutum* L. Theor Appl Genet 115: 675-682

Wang K, Song XL, Han ZG, Guo WZ, Yu JZ, Sun J, Pan JJ, Kohel RJ, Zhang TZ (2006b) Complete assignment of the chromosomes of *Gossypium hirsutum* L. by translocation and fluorescence in situ hybridization mapping. Theor Appl Genet 113: 73-80

Wang K, Yang ZJ, Shu CS, Hu J, Lin QY, Zhang WP, Guo WZ, Zhang TZ (2009b) Higher axial-resolution and sensitivity pachytene fluorescence in situ hybridization protocol in tetraploid cotton. Chromosome Res 17: 1041-1050

Wang PZ, Su L, Qin L, Hu BM, Guo WZ, Zhang TZ (2009a) Identification and molecular mapping of a Fusarium wilt resistant gene in upland cotton. Theor Appl Genet 119: 733-739

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

Wilkins MR, Pasquali C, Appel RD, Ou K, Golaz O, Sanchez JC, Yan JX, Gooley AA, Hughes G, Humphery-Smith, Williams KL, Hochstrasser DF (1996) From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis. Biotechnology14:61–65.

Wu AM, Lv SY, Liu JY (2007) Functional analysis of a cotton glucuronosyltransferase promoter in transgenic tobaccos. Cell Research 17: 174-183

Wu YT, Liu JY (2005) Molecular cloning and characterization of a cotton glucuronosyltranferase gene. J Plant Physiol 162: 573-582

Xiao YH, Li DM, Yin MH, Li XB, Zhang M, Wang YJ, Dong J, Zhao J, Luo M, Luo XY, Hou L, Hu L, Pei Y (2010) Gibberellin 20-oxidase promotes initiation and elongation of cotton fibers by regulating gibberellin synthesis. J Plant Physiol 167: 829-837

Xu WL, Wang XL, Wang H, Li XB (2007b) Molecular characterization and expression analysis of nine cotton *GhEF1A* genes encoding translation elongation factor 1A. Gene 389: 27-35

Xu Y, Li HB, Zhu YX (2007a) Molecular biological and biochemical studies reveal new pathways important for cotton fiber development. J Integr Plant Biol 49 (1): 69–74

Xue TT, Li XZ, Zhu W, Wu CA, Yang GD, Zheng CC (2009) Cotton metallothionein GhMT3a, a reactive oxygen species scavenger, increased tolerance against abiotic stress in transgenic tobacco and yeast. J Exp Bot 60(1):339–349

Yang C, Guo WZ, Li GY, Gao F, Lin SS, Zhang TZ (2008a) QTLs mapping for Verticillium wilt resistance at seedling and maturity stages in *Gossypium barbadense* L. Plant Science 174(3): 290-298

Yang CQ, Lu S, Mao YB, Wang LJ, Chen XY (2010) Characterization of two NADPH: cytochrome P450 reductases from cotton (*Gossypium hirsutum*). Phytochemistry 71: 27-35

Yang YW, Bian SM, Yao Y, Liu JY (2008c) Comparative proteomic analysis provides new insights into the fiber elongating process in cotton. J Proteome Res 7: 4623-4637

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

Yin JM, Guo WZ, Yang LM, Liu LW, Zhang TZ (2006) Physical mapping of the Rf_1 fertility-restoring gene to a 100 kb region in cotton. Theor Appl Genet 112: 1318-1325

Yu JW, Yu SX, Lu CR, Wang W, Fan SL, Song MZ, Lin ZX, Zhang XL, Zhang JF (2007) High-density linkage map of cultivated allotetraploid cotton based on SSR, TRAP, SRAP and AFLP markers. J Integr Plant Biol 49(5): 716-724

Zhang BH, Pan XP (2009) Expression of MicroRNAs in Cotton. Molecular Biotechnology 42:269-274

Zhang BH, Wang QL, Wang KB, Pan XP, Liu F, Guo TL, Cobb GP, Anderson TA (2007) Identification of cotton microRNAs and their targets. Gene 397:26-37

Zhang HB, Li Y, Wang BH, Chee PW (2008) Recent Advances in Cotton Genomics. International Journal of Plant Genomics 2008:742304

Zhang TZ, Yuan YL, Yu J, Guo WZ, Kohel RJ (2003) Molecular tagging of a major QTL for fiber strength in Upland cotton and its marker-assisted selection. Theor Appl Genet 106: 262-268

Zhang ZS, Xiao YH, Luo M, Luo M, Li XB, Luo XY, Hou L, Li DM, Pei Y (2005) Construction of a genetic linkage map and QTL analysis of fiber-related traits in upland cotton (*Gossypium hirsutum* L.). Euphytica 144:91–99

Zhao L, Cai CP, Zhang TZ, Guo WZ (2009) Fine mapping of the red plant gene R_1 in upland cotton (*Gossypium hirsutum*). Chin Sci Bull 54(9): 1529-1533 (in Chinese)

Zhao PM, Wang LL, Han LB, Wang J, Yao Y, Wang HY, Du XM, Luo YM, Xia GX (2010) Proteomic identification of differentially expressed proteins in the Ligon lintless mutant of upland cotton (*Gossypium hirsutum* L.). J Proteome Res 9: 1076-1087

Zhu LF, Tu LL, Zhang XL,Nie YC, Guo XP, Xia QZ (2005) Construction and analysis of SSH library of *Gossypium barbadense* upon infection with *Verticillium dahliae*. Acta Genetica Sinica 32:528-532

Zhu XX, Zhu YC, Ai NJ, Liu RZ, Zhang TZ (2009) Gene differential expression at seedling stage in four cotton combinations hybridized by CRI-12 and its pedigree-derived lines. Acta Agronomica Sinica 35(9):1637–1645 (in Chinese)

Zhu YQ, Xu KX, Luo B, Wang JW, Chen XY (2003) An ATP-binding cassette transporter *GhWBC1* from elongating cotton fibers. Plant Physiol 133: 580-588

Zuo KJ, Wang J, Wu W, Chai Y, Sun XF, Tang KX (2005) Identification and characterization of differentially expressed ESTs of *Gossypium barbadense* infected by *Verticillium dahliae* with suppression subtractive hybridization. Mol Biol 39:191-199.