

RESEARCH ARTICLE

Open Access



# Novel insights into water-deficit-responsive mRNAs and lncRNAs during fiber development in *Gossypium hirsutum*

Nan Wu<sup>†</sup>, Jun Yang<sup>†</sup>, Guoning Wang, Huifeng Ke, Yan Zhang, Zhengwen Liu, Zhiying Ma<sup>\*</sup> and Xingfen Wang<sup>\*</sup>

## Abstract

**Background:** The fiber yield and quality of cotton are greatly and periodically affected by water deficit. However, the molecular mechanism of the water deficit response in cotton fiber cells has not been fully elucidated.

**Results:** In this study, water deficit caused a significant reduction in fiber length, strength, and elongation rate but a dramatic increase in micronaire value. To explore genome-wide transcriptional changes, fibers from cotton plants subjected to water deficit (WD) and normal irrigation (NI) during fiber development were analyzed by transcriptome sequencing. Analysis showed that 3427 mRNAs and 1021 long noncoding RNAs (lncRNAs) from fibers were differentially expressed between WD and NI plants. The maximum number of differentially expressed genes (DEGs) and lncRNAs (DERs) was identified in fibers at the secondary cell wall biosynthesis stage, suggesting that this is a critical period in response to water deficit. Twelve genes in cotton fiber were differentially and persistently expressed at  $\geq$  five time points, suggesting that these genes are involved in both fiber development and the water-deficit response and could potentially be used in breeding to improve cotton resistance to drought stress. A total of 540 DEGs were predicted to be potentially regulated by DERs by analysis of coexpression and genomic colocation, accounting for approximately 15.76% of all DEGs. Four DERs, potentially acting as target mimics for microRNAs (miRNAs), indirectly regulated their corresponding DEGs in response to water deficit.

**Conclusions:** This work provides a comprehensive transcriptome analysis of fiber cells and a set of protein-coding genes and lncRNAs implicated in the cotton response to water deficit, significantly affecting fiber quality during the fiber development stage.

**Keywords:** *Gossypium hirsutum*, Fiber, mRNA, Long noncoding RNA (lncRNA), Water deficit

## Background

Cotton is the major source of natural fibers and the most important raw material for the textile industry. Cotton fiber development is generally defined into four distinct but overlapping stages, including fiber initiation [FI, from 2 days before anthesis to 3~6 days

post anthesis (DPA)], fiber elongation (FE, primary cell wall biosynthesis, 3~20 DPA), secondary cell wall biosynthesis (SCWB, 16~40 DPA), and maturation (40~50 DPA) [1, 2]. FI, marking the start of fiber growth, is a key stage that determines cotton yield. Many genes regulating FI have been identified in cotton, such as the MYB transcription factors *GhMYB25* [3] and *GhMYB25-like* [4], protodermal factor *GbPDF1* [5], jasmonate zim-domain protein *GhJAZ2* [6], auxin efflux carrier *GhPIN3a* [7], MYB-MIXTA-like transcription factors *GhMML4* and *GhMML3* [8]. FE has been suggested as a crucial stage for determining the

\*Correspondence: mzhy@hebau.edu.cn; cotton@hebau.edu.cn

<sup>†</sup>Nan Wu and Jun Yang contributed equally to this work.

State Key Laboratory of North China Crop Improvement and Regulation, North China Key Laboratory for Crop Germplasm Resources of Education Ministry, Hebei Agricultural University, Baoding 071001, China



final fiber quality. Some genes have been characterized to play roles in this stage. For example, six *Gh14-3-3* genes are predominantly expressed at the FE stage, and overexpression of these genes promotes the longitudinal growth of fission yeast, indicating that they might participate in the regulation of fiber elongation [9]. Two transcription factors (TFs), *GhHOX3* (homoeodomain-leucine zipper TF) and *GhDEL65* (basic helix-loop-helix TF), positively regulate cotton fiber elongation [10, 11]. *GhLTPG1* (glycosylphosphatidylinositol anchored lipid transport protein) is abundantly expressed in elongating cotton fibers, and knockdown of *GhLTPG1* results in shorter fibers with the repression of FE-related gene expression [12]. Cotton fiber quality traits, including strength, micronaire, and maturity, are mostly determined at the stages of SCWB and maturation. The biology and genes involved in these two stages are much less understood and studied [13], with only MYB transcription factors reported [14, 15].

High-throughput RNA-seq technology has been used to understand complex responses and for the functional exploration of protein-coding genes. Nevertheless, it was reported that a large portion of RNAs in eukaryotes, such as humans and Arabidopsis, do not encode proteins and are known as noncoding RNAs (ncRNAs) [16, 17]. ncRNAs over 200 nucleotides in length are named long ncRNAs (lncRNAs), which are further classified as long intergenic ncRNAs (lincRNAs), natural antisense transcripts (lncNATs), long intronic ncRNAs, and lncRNAs partially overlapping with protein-coding genes [18]. Thousands of lncRNAs have been identified in several plants, expanding our understanding of the plant transcriptome. In Arabidopsis, 6510 lncRNAs were identified, among which approximately five hundred showed inducible expression patterns upon exposure to abscisic acid (ABA) and drought [19]. In *Ricinus communis*, 5356 lncRNAs were cataloged and potentially involved in regulating the development of endosperm and embryos in castor bean seeds [20]. A total of 23,651 novel lncRNAs were identified in Tibetan wild barley, of which 535 lncRNAs were differentially expressed in response to drought stress [21]. In addition, as many as 59,110, 57,944 and 40,858 actively-expressed putative lncRNAs were identified from three wheat varieties, including Kiziltan, TR39477 and TTD-22 varieties, respectively [22]. With rapid advances in whole genome sequencing analyses of cotton, including *Gossypium raimondii*, *G. arboreum*, *G. hirsutum* and *G. barbadense*, thousands of lncRNAs have been identified and shown to have potential functions in cotton growth [23, 24], fiber development (initiation and elongation) [25, 26], and various stress responses, including drought [27], salt [28], phytopathogen *Verticillium*

*dahliae* infection [29, 30] and piercing-sucking pest *Aphis gossypii* attack [31].

Although lncRNAs have been identified in large numbers of plants and are believed to have crucial roles in development and stress responses, the functional roles and the mechanism underlying the lncRNAs is insufficient [32]. Novel functional networks will likely be defined by predicting and characterizing the interaction between lncRNAs and potential targets [33]. Emerging evidence suggests that plant lncRNAs have various mechanisms of action, mainly studied in Arabidopsis at present. lncRNA (*ASCO-lncRNA*) functions in Arabidopsis lateral root (LR) meristems by interacting with nuclear speckle RNA-binding protein (NSR), which is an alternative splicing regulator [34]. In other words, lncRNAs can hijack NSRs to affect their binding to mRNA targets. lncRNAs can bind with miRNAs by complementary sequences; thus, alterations in lncRNA abundance can modulate the action of miRNAs on downstream protein-coding genes. This mechanism of inhibition of miRNA activity was defined as target mimicry [35]. For example, lncRNA *IPS1* (induced by phosphate starvation 1) functions in regulating the phosphate starvation response in Arabidopsis by imperfect interaction with miRNA miR-399, which can guide the cleavage of *PHO2* [35, 36]. lncRNA *FRILAIR* (fruit ripening-related long intergenic RNA), a target mimic of miR397, can modulate the expression of *LAC11a* involved in strawberry fruit ripening [37]. In addition, the alteration of lncRNA expression can affect the dynamic chromatin topology, which determines the expression of neighboring genes. For instance, lincRNA *APOLO* transcription regulates the formation of a chromatin loop encompassing the promoter of its neighboring gene *PID*, a key regulator of polar auxin transport [38].

Cotton is mostly grown commercially in semiarid and arid environments. Fiber yield and quality are greatly and periodically affected by drought stress, and the severity of the problem may increase due to global climate change [39, 40]. Therefore, breeding cotton cultivars with higher yield and better fiber quality under drought conditions is becoming more urgent [41]. Some information has been available about cotton fiber development and drought resistance by the characterization of some important genes and analysis of transcriptome profiling, but the regulatory mechanism has not been absolutely elucidated to date [13, 41]. The Yellow River Basin (YRB) is one of the three major cotton production regions in China. In most years, rainfall was significantly reduced during the cotton flowering period in the YRB, which seriously affected the growth and development of cotton. To guarantee and increase cotton yield, supplementary irrigation just before flowering (SIF) is a widely

used farm operation to compensate for the lack of rainfall in the YRB. However, the effect of SIFs on cotton fiber development has not been investigated in detail thus far, especially on fiber quality, and the underlying mechanism for cotton fiber cells responding to non-SIFs (water deficit). The present study presents the first research on the effects of water deficit on cotton fiber development by genome-wide analysis of mRNAs and lncRNAs in fiber cells. In particular, some bifunctional lncRNAs preferentially expressed during fiber development and involved in the water-deficit response were identified. These results will deepen our understanding of the molecular mechanism underlying fiber development under drought stress, and provide clues to accelerate the development of novel cotton cultivars with improved yield potential, fiber quality, and adaptability to drought conditions.

## Methods

### Plant material and irrigation treatments

*G. hirsutum* cv. Nongdamian 13 (ND13) was planted in a rainout shelter at the Teaching Experimental Station of Hebei Agricultural University (38°49'N, 115°26'E), Baoding, China. The rows of plants were spaced 100 cm apart and 30 cm between plants within a row. When the first white bloom was observed, the normal irrigation cotton field (NI) received irrigation with 675 m<sup>3</sup>/ha of water, but the water deficit field (WD) did not. All other agronomic management practices were kept normal and uniform for NI and WD. The soil drought level was determined with the soil relative water content (SRWC) as previously described [42]. The cultivar ND13 was collected from Hebei Province in China by Hebei Agricultural University and approved by Hebei Provincial Crop Variety Certification Committee with an accession number G10072. All necessary permissions for planting and investigating this cultivar were obtained from Hebei Agricultural University and the National Medium-term Gene Bank of Cotton in China, and the collection and research of this cultivar have complied with the Convention on the Trade in Endangered Species of Wild Fauna and Flora.

### Fiber analysis

Seven replicates were taken for each treatment (NI and WD). For each replication, mature seed fibers were randomly sampled from 20 naturally-open bolls on the middle section of cotton plants. After drying, seed cotton was treated with a roller gin (MPSY-20, River Machinery Factory, Xinxiang, Henan, China) for separation of lint and seed, which were weighed to determine lint percentage (LP). Weight of 100 seeds was expressed as seed index (SI). The lint index (LI) was calculated by the formula:  $LI = (SI \times LP) / (1 - LP)$ . The fiber qualities were determined with an HVI-MF 100 instrument (User Technologies,

Inc., USTER, Switzerland) at the Supervision, Inspection and Testing Center of Cotton Quality, Ministry of Agriculture, Anyang, China. The data were analyzed ( $P$ -values < 0.05 was considered statistically significant) by unpaired  $t$ -test using GraphPad Prism 8.0.2 software (GraphPad Software, San Diego, USA).

### Fiber sample collection, RNA isolation, library construction and sequencing

Flowers were tagged on the day of flowering as 0 DPA. Ovules (0 and 5 DPA) and fibers (10, 15, 20, 25, 30 and 35 DPA) were collected, frozen immediately in liquid nitrogen and stored at  $-80^{\circ}\text{C}$ . Samples (ovules and fibers) from five independent plants within each treatment group served as one biological replication. Total RNA was isolated using the EASYspin Plant RNA Kit (Aidlab, Beijing, China). Libraries were constructed using the NEBNext<sup>®</sup> Ultra<sup>™</sup> Directional RNA Library Prep Kit for Illumina<sup>®</sup> (NEB, USA) according to the manufacturer's instructions. Strand-specific sequencing was performed on the Illumina HiSeq 4000 platform (paired-end 150-bp reads).

### mRNA and lncRNA identification and differential expression analysis

All raw data were processed by removing reads containing adapter or ploy-N and reads with low quality. The clean reads were aligned to the *G. hirsutum* TM-1 genome (NAU-NBI v1.1) [43] using TopHat v2.0.9 [44]. The mapped reads for each sample were assembled by Cufflinks v2.1.1 in a reference-based approach to identify mRNA transcripts (fragments per kilobase per million mapped reads, FPKM  $\geq 1$ ) [45]. Then, five steps were adopted to screen out lncRNAs from assembled transcripts: (1) transcripts with one exon, low expression, and low credibility were removed; (2) transcripts with length < 200 bp were eliminated; (3) transcripts that overlapped with annotated exons in the database were filtered out; (4) transcripts with FPKM < 0.5 were removed (however for transcripts with single exon, the threshold value was 2); (5) finally, transcripts with protein coding potential by the Coding Potential Calculator with NCBI eukaryotic protein database ( $E$ -value < 1e-10) and Pfam Scan (v1.3) (default parameters) were excluded [46, 47]. FPKMs of both mRNAs and lncRNAs in each sample were calculated using Cuffdiff (v2.1.1), which also provides statistical routines for determining differential expression using a model based on the negative binomial distribution [45]. Corrected  $P$ -value < 0.05 and the absolute value of  $\log_2(\text{FPKM}_{\text{WD}}/\text{FPKM}_{\text{NI}}) < 1$  were set as the threshold for significantly differential expression when processing the data with Cuffdiff.

### Real-time quantitative PCR (qPCR) analysis

Total RNA was reverse transcribed into cDNA using TransScript<sup>®</sup> One-Step gDNA Remover and cDNA Synthesis SuperMix (TransGen Biotech, China), according to the manufacturer's instructions. qPCR was conducted on a QuantStudio<sup>™</sup> 1 Real-Time PCR system (Thermo Fisher Scientific, USA) with TransStart Top Green qPCR SuperMix (+Dye I/+Dye II) (TransGen Biotech, China). The expression levels of mRNAs and lncRNAs were normalized by *GhUBQ7* (ubiquitin, Genbank accession No. DQ116441.1) using the  $2^{-\Delta\Delta CT}$  method [48]. The primers used in the study are listed in Table S1.

### Functional analysis of differentially expressed genes (DEGs) and lncRNAs (DERs)

The functions of lncRNAs were predicted according to the functional annotations of their potential target genes, which may be regulated by lncRNAs using two patterns: *cis*-acting (genomic colocation) and *trans*-acting (coexpression). The coding genes in the 100 kb up- or down-stream of lncRNAs were identified as colocalized genes. Coexpressed genes were predicted by the correlation in the expression between lncRNAs and coding genes (Pearson's correlation coefficient  $\geq 0.95$  or  $\leq -0.95$ ). KOBAS 3.0 was used to test the statistical enrichment of genes in GO and KEGG

pathways. A corrected *P*-value  $\leq 0.05$  was considered significantly enriched.

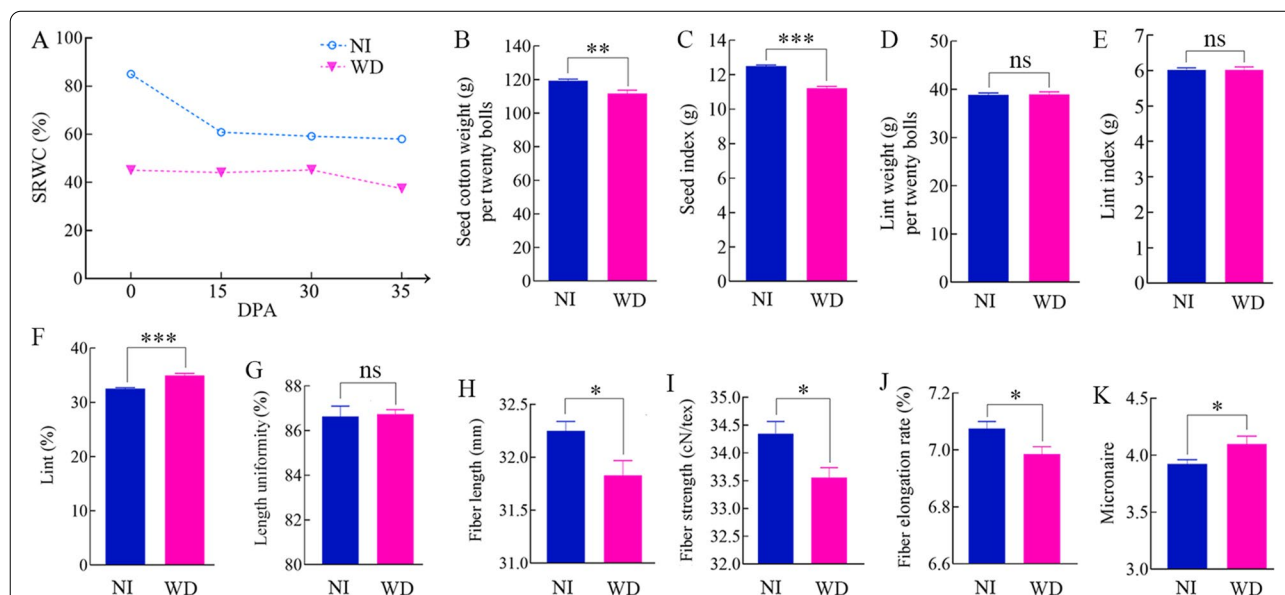
### Interaction prediction for lncRNA-miRNA-mRNA

The sequences of premiRNAs (precursor stem-loop molecules) and mature miRNAs in *G. hirsutum* were retrieved from the PNRD websites (<http://structuralbiology.cau.edu.cn/PNRD/index.php>) [49]. The targets (mRNAs and lncRNAs) of miRNAs were predicted using the psRNATarget online analysis tool (<http://plant.grn.noble.org/psRNATarget/analysis>) [50]. The potential interaction for lncRNA-miRNA-mRNA was constructed by (i) analyzing the same target miRNA for both lncRNA and mRNA and by (ii) evaluating the correlation in the expression between lncRNA and mRNA (Pearson's correlation coefficient  $\geq 0.95$  or  $\leq -0.95$ ).

## Results

### Water deficit reduces cotton fiber qualities

Two cotton fields received different treatments, NI and WD, at the beginning of the flowering stage. The SRWC in the NI-field (85.06%) was nearly twice as high as that in the WD-field (45.09%) at 0 DPA. During the following days, the SRWC in WD (37.47~45.17%) remained lower than that in NI (58.02~60.91%) (Fig. 1A) until cotton boll maturation began at 35 DPA. After another 30 days, fully mature cotton fibers were collected for weighing and quality determination. The seed cotton weight of the WD-group was significantly lower than that of the NI-group



**Fig. 1** Water deficit caused a significant reduction in fiber quality. **A** SRWC in the NI-field and WD-field at 0, 15, 30 and 35 DPA. Seed cotton weight (**B**), seed index (**C**), lint weight (**D**), lint index (**E**), lint (**F**), length uniformity (**G**), fiber length (**H**), fiber strength (**I**), fiber elongation rate (**J**), and micronaire (**K**), for NI-treatments were compared with WD-treatments. Data represent the mean  $\pm$  SE of seven biological replicates. Fibers for each replication were sampled from 20 naturally-open bolls on the middle section of cotton plants. Confidence levels were tested by unpaired *t*-test (\*,  $P < 0.05$ ; ns, not significant)

(Fig. 1B). The SI of the WD-group was also significantly lower than that of the NI-group (Fig. 1C), but there was no significant difference in lint weight (Fig. 1D) or LI (Fig. 1E) between the two groups. These results indicated that water deficit affected cotton seed weight but did not significantly affect fiber weight. Thus, water deficit-induced seed weight reduction resulted in a significant increase in LP (Fig. 1F). The uniformity ratio of fiber was not significantly reduced in WD compared with NI (Fig. 1G). However, water deficit caused a significant reduction in fiber length (Fig. 1H), strength (Fig. 1I) and elongation rate (Fig. 1J), and a dramatic increase in micronaire value (Fig. 1K). These results indicated that water deficit during cotton fiber development could lead to lower fiber qualities.

### Overview of fiber transcriptomes

To explore the transcriptional regulation of cotton fiber development under water deficit, the fiber transcriptomes of *G. hirsutum* from FI to SCWB were sequenced using RNA-seq. cDNA libraries from eight time points were constructed, which included 0, 5, 10, 15, 20, 25, 30 and 35 DPA. The flowering day was designed as 0 DPA, which represents FI. Both 10 and 15 DPA represent FE. To evaluate the SCWB, three time points including 25, 30 and 35 DPA were chosen. In addition, 5 and 20 DPA represent fiber developmental transitions (FDT1 and FDT2), which are stages from FI to FE and from FE to SCWB, respectively (Fig. 2A). Approximately 3.22 billion clean reads were screened out from 3.26 billion raw reads, varying from 83 to 116 million reads per library. The mapping rates of each library to the reference genome of *G. hirsutum* TM-1 ranged from 80.72% to 89.27% (Table S2). The correlation coefficients for two biological replicates at each time point were all above 0.85 (Table S3), indicating that the RNA-seq data have high reproducibility.

In total, 47,095 mRNAs (putative protein-coding genes) were identified with FPKM  $\geq 1$  and annotated according to the reference genome of *G. hirsutum* (Table S4). After multistep filtering, 13,051 high-confidence lincRNAs were identified, including 11,683 lincRNAs (89.52%) and 1368 lincNATs (10.48%) (Fig. 2B, Table S5). The transcript length of lincNATs (mean = 1883 nt) was significantly longer than that of lincRNAs (mean = 1327 nt) and mRNAs (mean = 1307 nt) (Fig. 2C). Most lincRNAs and lincNATs contained fewer than 6 exons, while mRNAs contained various numbers of exons (Fig. 2D). The overall

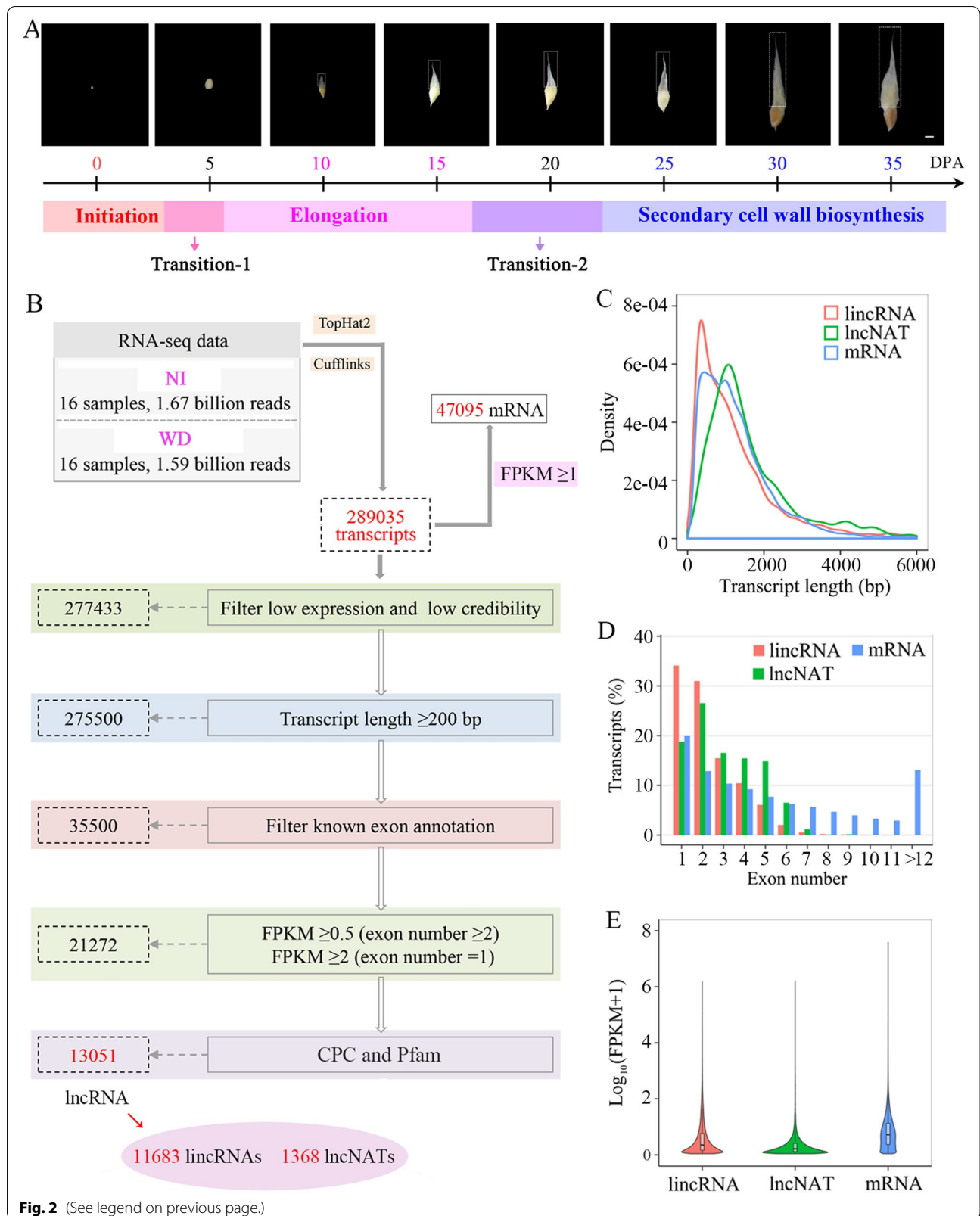
expression levels of both lincRNAs and lincNATs were lower than those of mRNAs (Fig. 2E). To validate the reliability of the transcriptome, 10 mRNAs and 10 lincRNAs were randomly selected for expression analysis by qPCR. For most mRNAs and lincRNAs, the linear regression analysis revealed a positive correlation between the transcriptome data and the results from qPCR with  $r$ -values (Fig. 3), suggesting the high quality of transcriptomes. Only one lincRNA, LNC009310, showed a relatively low correlation between transcriptome and qPCR analysis ( $r = 0.55$ ), likely due to the low expression of lincRNAs [51].

### Identification of differentially expressed genes (DEGs) in cotton fibers between NI and WD

A total of 3427 DEGs with at least a twofold expression change ( $FPKM_{WD}/FPKM_{NI}$ , corrected  $P$ -value < 0.05) were identified (Fig. 4, Table S6). During the stages of FI, SCWB, and FDTs, most DEGs were downregulated, whereas approximately 87% of DEGs were upregulated at the FE stage. At the two FDT stages, fewer DEGs were identified compared with the other three stages of fiber development. The maximum number of DEGs was identified at the SCWB stage with 2265, among which the greatest number of genes were differentially expressed at 30 DPA. Furthermore, the expression specificity of these DEGs at different stages of cotton fiber development was observed. Many genes were only differentially expressed at one time point, suggesting that they have time-specific expression in fibers under the stress of water deficit. For example, 821 DEGs were shown to be differentially expressed only at 30 DPA. No DEGs appeared at all 8 time points, but 12 genes were differentially expressed at  $\geq 5$  time points, including *ADH* (alcohol dehydrogenase), *MIOX* (*myo*-inositol oxygenase), *TK* (thymidine kinase), *PS* (phosphate starvation-induced gene), *PIP* (plasma-membrane intrinsic protein), *PAP* (purple acid phosphatase), *SPX* (SYG1-Pho81-XPR1 domain-containing protein), *NAM* (no apical meristem), *NCED* (9-cis-epoxycarotenoid dioxygenase), *UMAMIT* (usually multiple acids move in and out transporter) and *PEPC* (phosphoethanolamine/phosphocholine phosphatase), indicating that they maintain an intense response to the stress of water deficit.

(See figure on next page.)

**Fig. 2** Identification and characterization of mRNAs and lincRNAs in *G. hirsutum* fibers. **A** Representative images of individual seeds with attached fibers from 0 DPA to 35 DPA. Cotton fibers undergo three major sequential and overlapping developmental stages before maturity, including initiation, elongation and secondary cell wall biosynthesis. Transition-1 and -2 are two fiber developmental transition stages, which are from initiation to elongation and from elongation to secondary cell wall biosynthesis, respectively. The scale bars in all panels are 0.5 cm. **B** The pipeline of mRNAs and lincRNAs identification. **C** Length density distributions of lincRNAs, lincNATs and mRNAs. **D** Exon number per transcript of lincRNAs, lincNATs and mRNAs. **E** FPKM distributions of lincRNAs, lincNATs and mRNAs



### Functional classification of DEGs

To further characterize the functional consequences of gene expression changes in cotton fiber cells associated with water deficit, pathway enrichment analyses for DEGs were performed using the KEGG database (Fig. 5). At the FI stage, DEGs were significantly enriched in 5 pathways, especially “plant hormone signal transduction” and “photosynthesis-antenna proteins”. At the FE stage, DEGs were significantly enriched in 11 pathways, mainly “plant hormone signal transduction” and “phenylpropanoid biosynthesis”. At the SCWB stage, “plant hormone signal transduction” was no longer enriched, and 17 pathways related to metabolism and 1 pathway related to ABC transporters were significantly enriched. The pathways of “lipid metabolism”, “energy metabolism”, “amino acid metabolism” and “phenylpropanoid biosynthesis” are possibly involved in the further development of fiber and the cell wall. Only three pathways related to genetic information processing were enriched at FDT1, and no pathway was enriched at FDT2. In addition, almost all DEGs involved in enriched pathways for FI, FDT1 and SCWB were significantly downregulated, suggesting that these biological processes were suppressed under the stress of water deficit. Instead, almost all DEGs involved in pathways at the FE stage were significantly upregulated, suggesting that these biological processes were activated.

### Different expression profiles of lncRNAs in cotton fibers under water deficit

Under water deficit stress, a total of 1021 lncRNAs were differentially expressed (DERs) (Table S7), of which the majority were significantly downregulated (Fig. 6). In addition, the majority of DERs were only expressed at a specific time point (Fig. 6), showing similar expression characteristics to DEGs. Up to 700 DERs were identified at the SCWB stage, and many lncRNAs were specifically and differentially expressed at 25 DPA, as many as 315 in total, suggesting that lncRNAs mainly play a role in SCWB for fibers in response to water deficit.

### DEGs expression were potentially regulated by DERs

By analysis of coexpression and genomic colocation a total of 540 DEGs were predicted to be potentially

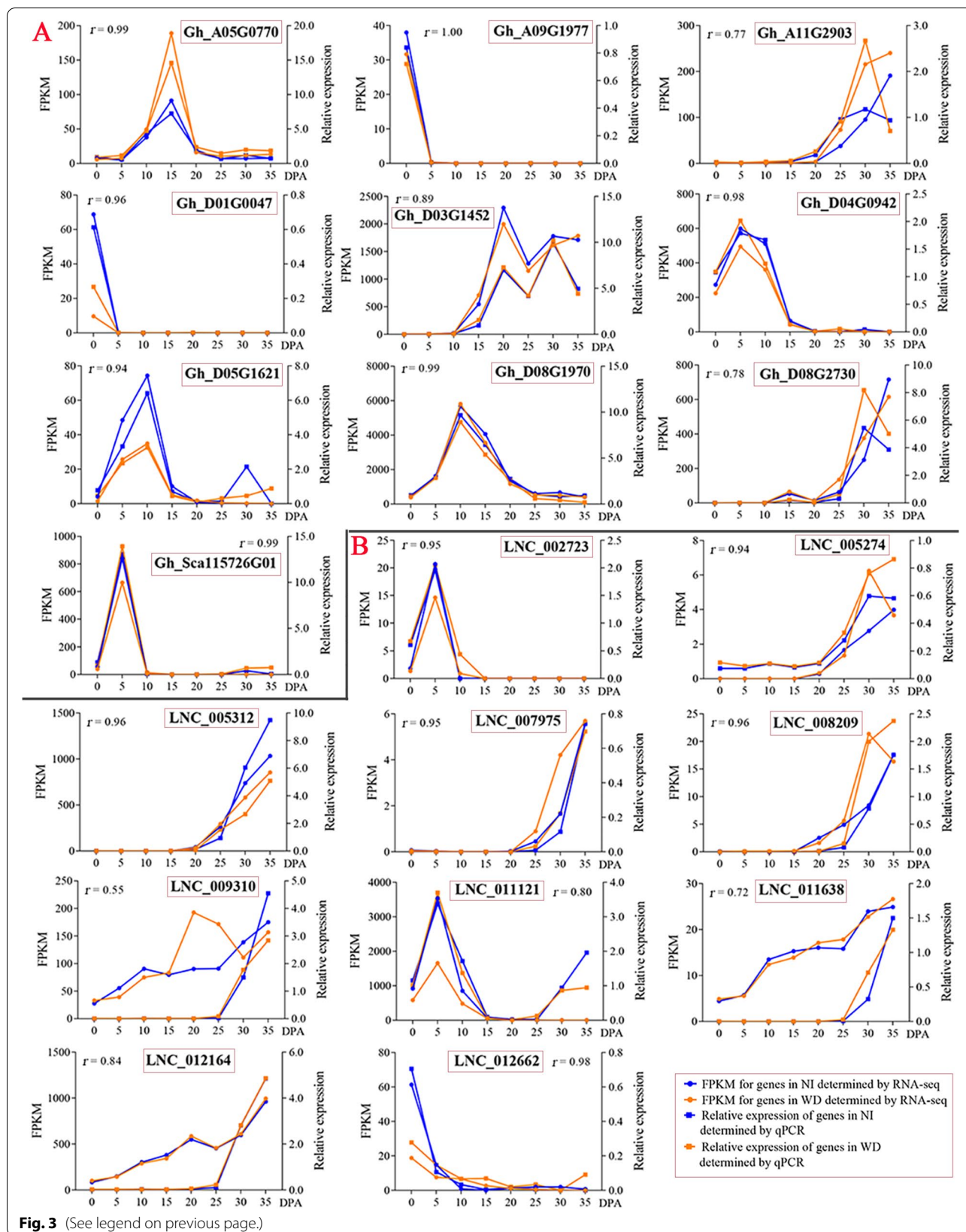
regulated by DERs (Table S8), accounting for approximately 15.76% of all DEGs. As shown in Fig. 7, the largest number of DEGs regulated by DERs (DEGs-R) was found at the SCWB stage (25, 30 and 35 DPA). However, according to the proportion calculation, the FI stage (0 DPA) had a larger proportion of DEGs-R (25.44%). DEGs-R were used for further GO enrichment analysis (Table S9). GO terms “cell wall organization or biogenesis” and “cell wall macromolecule metabolic process” in the biological process category were enriched, suggesting that lncRNAs targeting mRNA mainly regulate cell wall development. GO terms of “DNA packaging complex”, “protein-DNA complex”, “nucleosome”, “chromatin” and “chromosomal part” in the cellular component category were enriched, indicating that lncRNAs play regulatory roles in the nucleus. Only one GO term in the molecular function category was enriched (“Hydrolase activity”), showing that lncRNAs mainly function by regulating hydrolases. In addition, four pairs of regulatory relationships between DEGs and DERs mediated by miRNAs were predicted (Fig. 8), including LNC\_006412::ghr-miR482c::Gh\_A10G1972, LNC\_008673::ghr-n68::Gh\_D06G2174, LNC\_010115::ghr-miR482h/ghr-miR6118\*::Gh\_A07G2348/Gh\_D07G0162, and LNC\_004724::ghr-miR403::Gh\_A07G2019. These DERs may potentially combine with miRNAs, which also possibly interact with DEGs, using sequence complementation containing several mismatches.

### Discussion

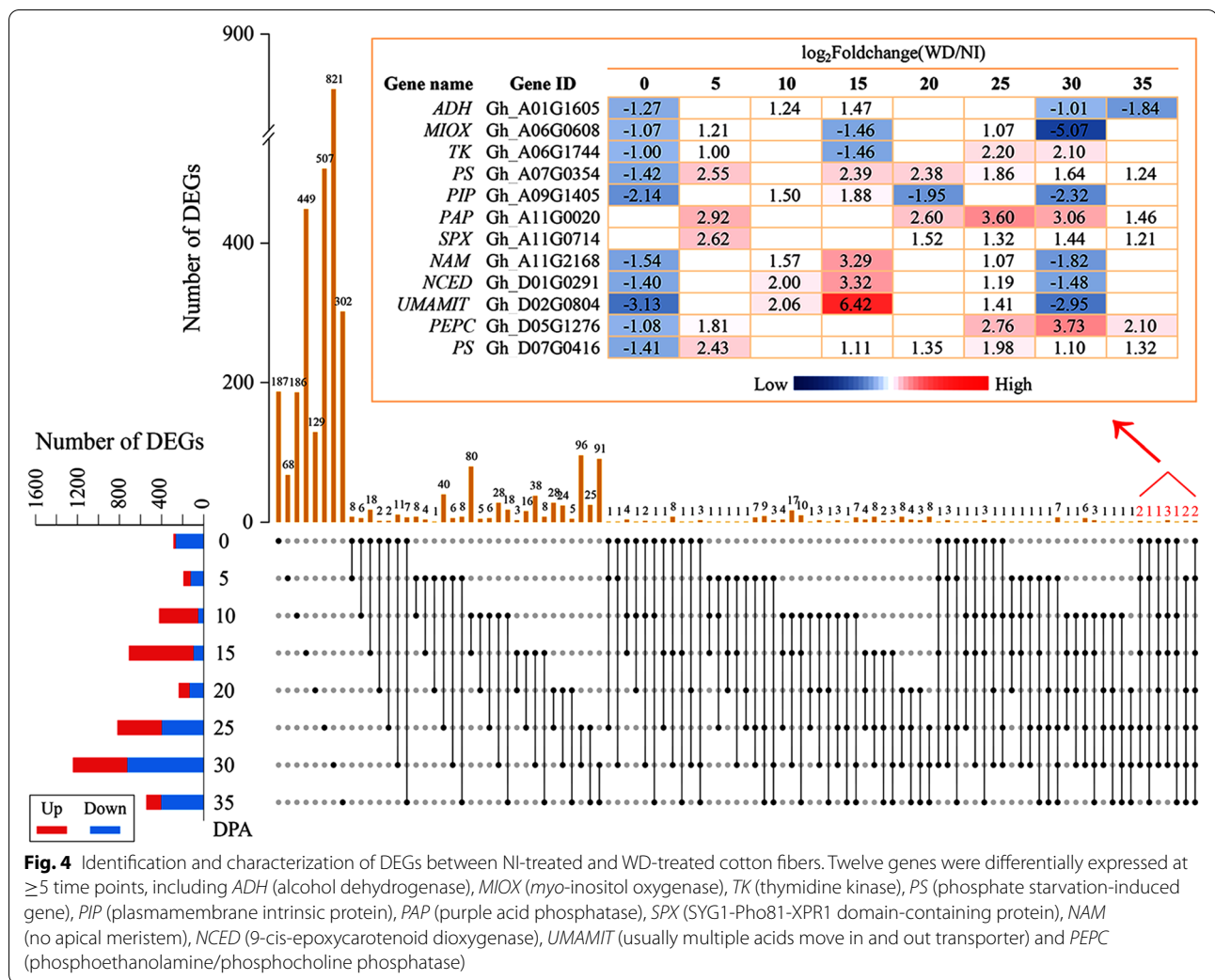
Water deficiency is one of the most impactful stresses worldwide and has long been prevalent in many countries, leading to reductions in cotton productivity and fiber quality. This negative effect varies depending on cotton growth stages and the intensity of water deficit [52]. SIF is always applied as an effective measure for guaranteeing and increasing cotton yield in the YRB regions of China. Here, our results strongly support the importance of SIF in cotton fiber development. From 0 DPA (bloom) to 35 DPA (fiber maturity), the SRWC of cotton fields that received SIF was significantly higher than that of cotton fields without SIF (Fig. 1A). Therefore, the cotton plants grown in the field without SIF were considered to be subjected to water deficit stress throughout cotton

(See figure on next page.)

**Fig. 3** Confirmation of the expression patterns of mRNAs (A) and lncRNAs (B) using qPCR. Ten mRNAs and ten lncRNAs were randomly selected for expression analysis during the fiber developmental stages of *G. hirsutum* ND13 treated with NI and WD. The correlation of relative expression for mRNAs and lncRNAs measured by RNA-seq and qPCR was estimated with *r*-values. *UBQ7* was used as the reference gene. Gene (mRNA) IDs are shown in the genome of *G. hirsutum* TM-1 (NAU-NBI v1.1), including Gh\_A05G0770 (17.3 kDa class I heat shock protein), Gh\_A09G1977 (1-aminocyclopropane-1-carboxylate oxidase homolog 1), Gh\_A11G2903 (ABC transporter G family member 2), Gh\_D01G0047 (Protein RADIALIS-like 6), Gh\_D03G1452 (Tubulin beta-7 chain), Gh\_D04G0942 (No annotation), Gh\_D05G1621 (No annotation), Gh\_D08G1970 (Probable aquaporin PIP1-2), Gh\_D08G2730 (Bidirectional sugar transporter SWEET15), and Gh\_Sca115726G01 (Aspartic proteinase nepenthesin 1)



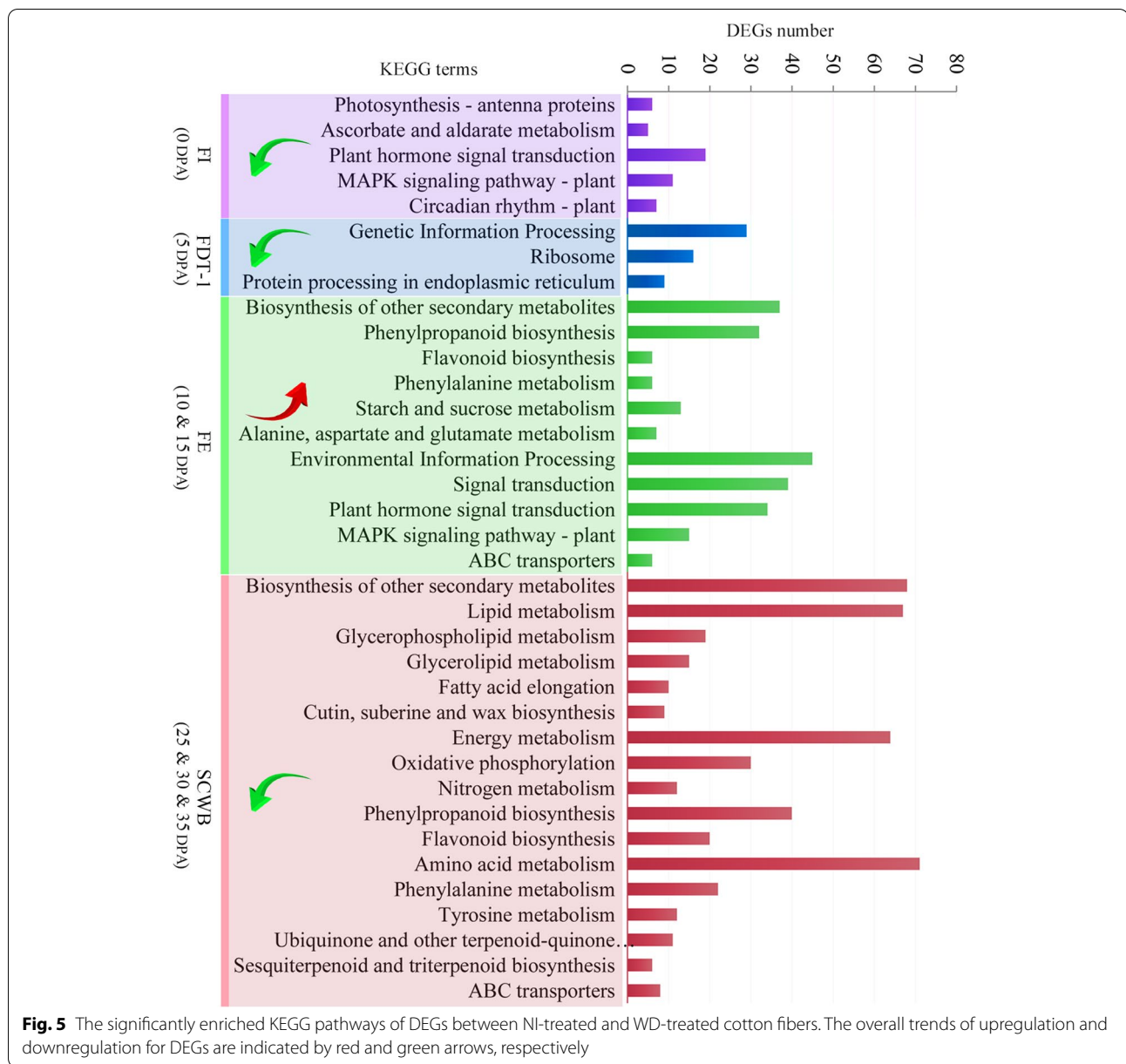




fiber development. As expected, water deficit caused a significant reduction in seed cotton yield (Fig. 1B and C). However, lint yield did not significantly decrease because of the water deficit (Fig. 1D and E). As a result, there was a significant increase in the percentage of lint (Fig. 1F). Therefore, it can be concluded that the loss of seed cotton yield is mainly due to the reduction in seed weight, which is caused by non-SIF. Furthermore, the quality of fibers from WD was compared with that from NI. Water deficit caused a significant reduction in fiber length (Fig. 1H), fiber strength (Fig. 1I), elongation rate (Fig. 1J) but a dramatic increase in micronaire (Fig. 1K), suggesting that fiber development is very sensitive to water deficit. In addition, micronaire is an indicator of air permeability and universally used for assessing fiber maturity (degree of secondary cell-wall development) and fineness [53]. Here, fully mature cotton fibers (65 DPA) were collected for weighing and quality determination. Thus, the micronaire value mainly represents the thickness of the fiber.

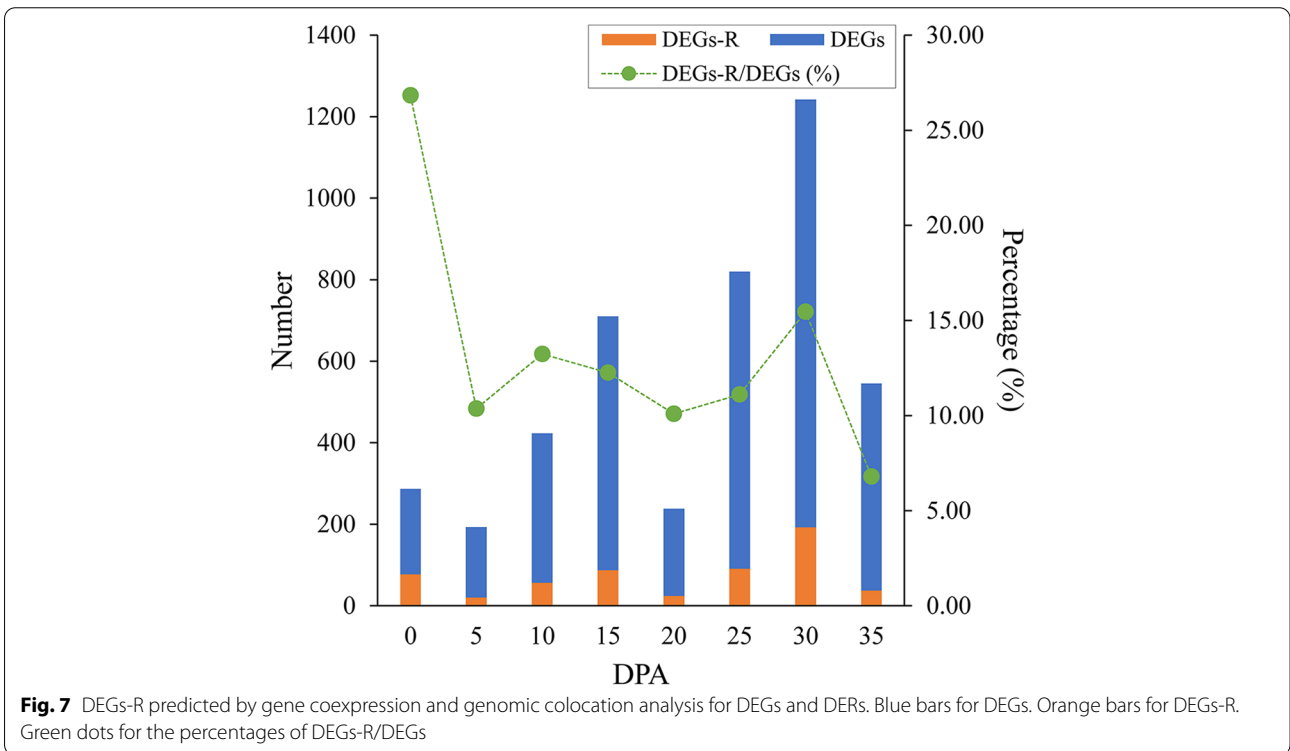
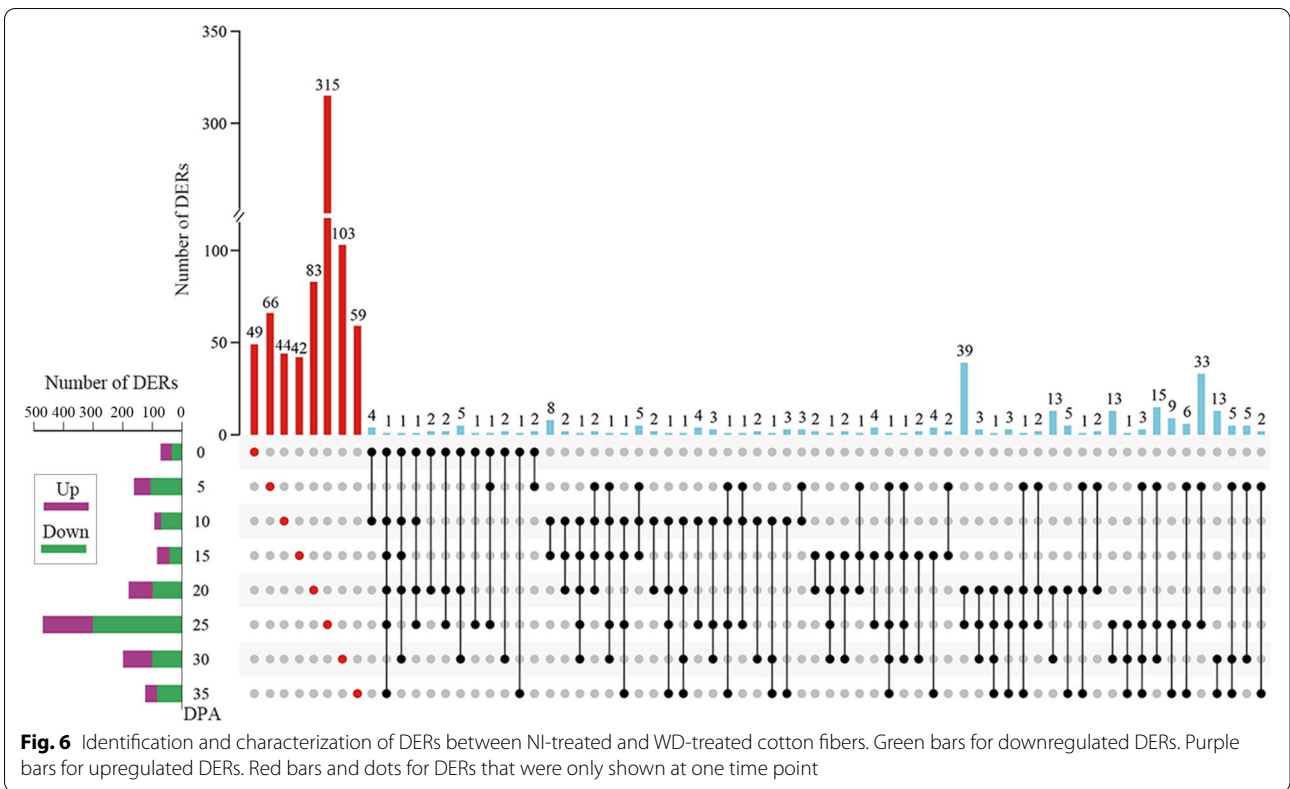
Additionally, water deficit usually induces the thickening of cell walls, which is an important adaptation to increase plant tolerance to water loss [54]. Therefore, it can be inferred that water deficit makes cotton fibers thicker. Fibers with micronaire values that are too high or too low are undesirable from the point of view of spinning and yarn evenness. Micronaires have been shown to increase or decrease with irrigation changes [55, 56]. Here, water deficit significantly increased the micronaire value of ND13. Therefore, the expected micronaires might be obtained by properly controlling SIF in the future to meet the demand of the cotton industry.

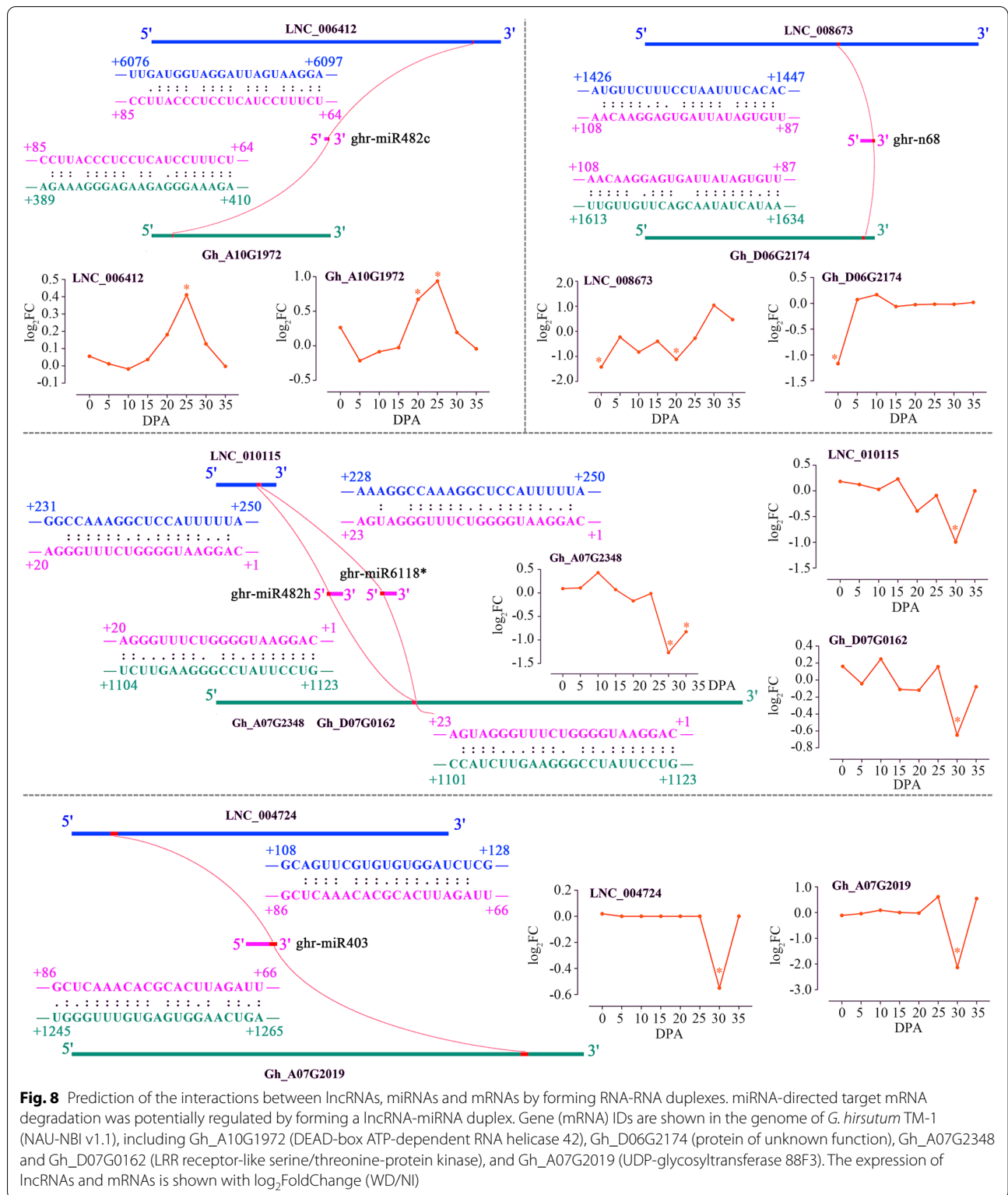
To further our knowledge of the molecular mechanisms underlying fiber cells response to water deficit, a genome-wide identification and characterization of water deficiency-responsive genes and lncRNAs was carried out in this study. At the FI and FE stages, the pathway “plant hormone signal transduction” was enriched (Fig. 5), suggesting that water deficit affected the



regulatory networks controlled by various hormones that are necessary for fiber initiation and elongation. According to the number of DEGs, auxin, ethylene and ABA are the most important signal regulatory molecules for cotton fibers in response to water deficit. This is consistent with previous reports that these hormones are involved not only in fiber development but also in plant drought

resistance [57–59]. At the FE stage, up to 32 DEGs were involved in “phenylpropanoid biosynthesis”, which participates in the biosynthesis of many plant cell wall phenolic products, such as lignins, flavonoids, suberins, and cutins [60, 61]. The significantly upregulated expression of these DEGs may affect fiber cell expansion and prematurely end cell elongation, ultimately leading to a significant





reduction in fiber length. However, at the SCWB stage, DEGs involved in “phenylpropanoid biosynthesis” were changed to be significantly downregulated, which may

reduce cell wall thickening and lignin deposition [62]. Although the content of lignin and lignin-like phenolics is minor in cotton fibers, recent data suggest that these

ingredients are strongly associated with fiber strength and elongation [62–64].

In addition, the expression of most DEGs showed temporal specificity; that is, they were expressed only at one point in time. However, 12 genes were differentially expressed at more than 5 time points (Fig. 4), suggesting that they can intensely and persistently respond to water deficit. MIOX is known to balance the concentrations of *myo*-inositol and UDP-GlcA (UDP-glucuronic acid). *Myo*-inositol plays an important role in drought tolerance by scavenging reactive oxygen species, decreasing the loss of chlorophyll for photosynthesis, and improving antioxidant enzyme activity [65, 66]. UDP-GlcA is the precursor for UDP-xylose, which is a critical component of cell wall polysaccharides, such as pectin and hemicellulose [67]. PIPs are major facilitators that conduct water and/or other molecules across cell membranes. Thus, PIPs are usually responsive to drought stress and play pivotal roles in plant drought resistance by regulating the transcellular transport of water [68]. Meanwhile, PIPs can selectively form primary aquaporin isoforms to meet the requirements for rapid elongation of fibers [69]. Thus, *MIOX* and *PIP* in cotton fibers were differentially and persistently expressed, suggesting that they are difunctional genes involved in both fiber development and drought resistance, and can be potentially used in breeding to improve cotton resistance to water deficit.

Previously, thousands of lncRNAs have been identified and proposed to have functions in fiber development [23, 25], resistance to *V. dahliae* [29, 30], response to drought [27] and salt [28]. Here, our study provides the first comprehensive identification of lncRNAs in fiber cells of *G. hirsutum* under water deficit conditions. Up to 700 DERs (approximately 68.56% of the total DERs) were identified at the SCWB stage (Fig. 6), which is key for determining fiber length and strength. In addition, more than 300 DERs were specifically expressed at 25 DPA. These results suggest that the expression of lncRNAs in fibers changes in response to water deficit and differs significantly depending on fiber development. Understanding on the mechanisms of lncRNA action in plants remains limited and a major challenge [32]. lncRNAs have no discernable protein coding potential, or can encode only small peptides, but often result in functional RNAs involved in a wide range of molecular processes including but not limited to all steps of gene expression involving nucleic acids from chromatin modifications to translation [70]. Cotton lincRNA DAN1, a well-known example for transcriptional regulation, can bind DNA sequences containing AAAG motifs hence silencing of DAN1 increased cotton drought tolerance by regulating auxin responsive genes with AAAG motifs

[71]. Target mimicry is a generally accepted and prediction available post-transcriptional regulating mechanism by which lncRNA affect mRNA expression by regulating miRNA activity [35]. lncRNAs are called miRNA “sponges” that sequester miRNAs with imperfect base complementarity. In cotton, lncRNA354 is negatively related to salt tolerance by regulating ARF genes through miR160b [72]. Here, with WD, 540 DEGs were predicted to be potentially regulated by DERs by analysis of coexpression and genomic colocalization with mRNA. Furthermore, four pairs of regulatory relationships between DEGs and DERs mediated by miRNAs were predicted (Fig. 8). These four lncRNAs potentially interact with a miRNA by forming an lncRNA-miRNA duplex, which functions as RNA interference (RNAi), and as a result, miRNA-targeted mRNA is normally translated into a functional protein [32]. MiR403 has been reported to be involved in plant drought [73–75], heat, salt and cadmium stress response in a tissue associated manner [73]. Additionally, miR403 has an antiviral role by controlling the expression of AGO2 (Argonaute 2) [76]. In our study, the putative target gene of ghr-miR403, *UGT88F3* (UDP-glycosyltransferase 88F3, Gh\_A07G2019) is likely involved in fiber development and osmotic stress [77]. MiR482 is an ancient microRNA family present in all land plants. In tomato, overexpression of miR482c induced enhanced susceptibility to late blight while knock out miR482c and miR482b simultaneously enhanced resistance to late blight and the effect was stronger than silencing miR482b alone, possibly by regulating expression levels of genes encoding for proteins with nucleotide binding sites and leucine-rich repeat (NBS-LRR) domains and ROS levels [78, 79]. As in tomato, cotton plants are able to induce expression of NBS-LRR defence genes by suppressing the miR482-mediated gene silencing pathway upon fungal pathogen attack [80]. Particularly, miR482c might participate in cotton growth and abiotic stresses including drought stress response as a potential regulator [81]. As sequence variation induced target and mechanism variation, the function of members in miR482 family differ possibly from each other [78, 82]. *RH42* (DEAD-box ATP-dependent RNA helicase 42, Gh\_A10G1972) plays an essential role in DNA and RNA metabolism such as transcription, replication and repair [83]; *ERECTA* (LRR receptor-like serine/threonine-protein kinase, Gh\_A07G2348 and Gh\_D07G0162) has been shown to regulate plant flowering and transpiration efficiency in part via effects on epidermal cell expansion, cell-cell communication, and stomatal density [84, 85]. Gh\_D06G2174 with a DUF616 domain (protein of unknown function) has not been described and annotated in the genomic database of *G. hirsutum*. Although DEGs potentially

regulated by lncRNAs and miRNAs have been successfully found here, more experiments are needed to further confirm the interactions between them.

## Conclusions

Water deficit during cotton fiber development caused a significant reduction in cotton seed yield and fiber quality. Through the identification and functional classification of DEGs and DERs in cotton fibers between NI and WD, a valuable platform for revealing the molecular mechanism of cotton against water deficit was provided. In addition, potential functions for some lncRNAs regulating mRNA transcription were predicted, which provides valuable information to further characterize their functions.

## Abbreviations

ABA: Abscisic acid; DEGs: Differentially expressed genes; DEGs-R: DEGs regulated by DERs; DERs: Differentially expressed lncRNAs; DPA: Days post anthesis; FDT: Fiber developmental transition; FE: Fiber elongation; FI: Fiber initiation; FPKM: Fragments per kilobase per million mapped reads; lincRNAs: Long intergenic noncoding RNAs; lincNATs: Long noncoding natural antisense transcripts; lncRNAs: Long noncoding RNAs; LP: Lint percentage; NI: Normal irrigation; qPCR: Real-time quantitative PCR; RNAi: RNA interference; SCWB: Secondary cell wall biosynthesis; SI: Seed index; SIF: Supplementary irrigation just before flowering; SRWC: Soil relative water content; YRB: The Yellow River Basin; WD: Water deficit.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12870-021-03382-y>.

**Additional file 1: Table S1.** Primers used for qPCR.

**Additional file 2: Table S2.** *G. hirsutum* fiber transcriptomes assembly statistics.

**Additional file 3: Table S3.** The correlation ( $R^2$ ) of gene expression ( $\log_{10}(\text{FPKM} + 1)$ ) between two biological replicates.

**Additional file 4: Table S4.** FPKMs of 47,095 mRNAs in cotton fibers under NI and WD.

**Additional file 5: Table S5.** FPKMs of 13,051 lncRNAs in cotton fibers under NI and WD.

**Additional file 6: Table S6.** List of DEGs in cotton fibers between NI and WD.

**Additional file 7: Table S7.** List of DERs in cotton fibers between NI and WD.

**Additional file 8: Table S8.** List of DEGs potentially regulated by DERs.

**Additional file 9: Table S9.** GO enrichment analysis of DEGs-R.

## Acknowledgements

All authors are grateful to the laboratory members for help, advice and discussion.

## Authors' contributions

NW and JY carried out the experiment. GW and ZL contributed to the interpretation of the results. HK and YZ contributed to sample preparation. JY and NW wrote the manuscript with support from XW and ZM. XW and ZM

conceived the original idea and supervised the project. All authors have read and approved the manuscript.

## Funding

This work was supported by grants from the National Natural Science Foundation of China (31971982) and the Top Talent Fund of Hebei Province. The funding bodies provided the financial support to this research, including experimental design and implementation, sampling and data analysis. No funder played the role in data collection and analysis and writing the manuscript.

## Availability of data and materials

The data generated or analyzed during the current study are included in this published article and its supplemental data files. The RNA-seq data are available in the Genome Sequence Archive (<https://ngdc.cnpc.ac.cn>; accession number: CRA005441) or from the corresponding author on reasonable request.

## Declarations

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Competing interests

The authors declare that they have no conflict of interest.

Received: 12 June 2021 Accepted: 5 December 2021

Published online: 03 January 2022

## References

1. Qin YM, Zhu YX. How cotton fibers elongate: a tale of linear cell-growth mode. *Curr Opin Plant Biol.* 2011;14(1):106–11.
2. Huang G, Huang JQ, Chen XY, Zhu YX. Recent advances and future perspectives in cotton research. *Annu Rev Plant Biol.* 2021;72:437–62.
3. Machado A, Wu YR, Yang YM, Llewellyn DJ, Dennis ES. The MYB transcription factor *GhMYB25* regulates early fibre and trichome development. *Plant J.* 2009;59(1):52–62.
4. Walford SA, Wu YR, Llewellyn DJ, Dennis ES. GhMYB25-like: a key factor in early cotton fibre development. *Plant J.* 2011;65(5):785–97.
5. Deng F, Tu L, Tan J, Li Y, Nie Y, Zhang X. GbPDF1 is involved in cotton fiber initiation via the core *cis*-element HDZIP2ATATHB2. *Plant Physiol.* 2012;158(2):890–904.
6. Hu H, He X, Tu L, Zhu L, Zhu S, Ge Z, et al. GhJAZ2 negatively regulates cotton fiber initiation by interacting with the R2R3-MYB transcription factor GhMYB25-like. *Plant J.* 2016;88(6):921–35.
7. Zhang M, Zeng JY, Long H, Xiao YH, Yan XY, Pei Y. Auxin regulates cotton fiber initiation via GhPIN-mediated auxin transport. *Plant Cell Physiol.* 2017;58(2):385–97.
8. Wu HT, Tian Y, Wan Q, Fang L, Guan XY, Chen JD, et al. Genetics and evolution of MIXTA genes regulating cotton lint fiber development. *New Phytol.* 2018;217(2):883–95.
9. Zhang ZT, Zhou Y, Li Y, Shao SQ, Li BY, Shi HY, et al. Interactome analysis of the six cotton 14-3-3s that are preferentially expressed in fibres and involved in cell elongation. *J Exp Bot.* 2010;61(12):3331–44.
10. Shan CM, Shangguan XX, Zhao B, Zhang XF, Chao LM, Yang CQ, et al. Control of cotton fibre elongation by a homeodomain transcription factor *GhHOX3*. *Nat Commun.* 2014;5:5519.
11. Shangguan XX, Yang CQ, Zhang XF, Wang LJ. Functional characterization of a basic helix-loop-helix (bHLH) transcription factor GhDEL65 from cotton (*Gossypium hirsutum*). *Physiol Plant.* 2016;158(2):200–12.
12. Deng T, Yao H, Wang J, Wang J, Xue H, Zuo K. GhLTPG1, a cotton GPI-anchored lipid transfer protein, regulates the transport of phosphatidylinositol monophosphates and cotton fiber elongation. *Sci Rep.* 2016;6:26829.

13. Fang DD, Naoumkina M, Kim HJ. Unraveling cotton fiber development using fiber mutants in the post-genomic era. *Crop Sci*. 2018;58(6):2214–28.
14. Huang JF, Guo YJ, Sun QW, Zeng W, Li J, Li XB, et al. Genome-wide identification of R2R3-MYB transcription factors regulating secondary cell wall thickening in cotton fiber development. *Plant Cell Physiol*. 2019;60(3):687–701.
15. Sun X, Gong SY, Nie XY, Li Y, Li W, Huang GQ, et al. A R2R3-MYB transcription factor that is specifically expressed in cotton (*Gossypium hirsutum*) fibers affects secondary cell wall biosynthesis and deposition in transgenic *Arabidopsis*. *Physiol Plant*. 2015;154(3):420–32.
16. Chekanova JA, Gregory BD, Reverdatto SV, Chen H, Kumar R, Hooker T, et al. Genome-wide high-resolution mapping of exosome substrates reveals hidden features in the *Arabidopsis* transcriptome. *Cell*. 2007;131(7):1340–53.
17. Kapranov P, Cheng J, Dike S, Nix DA, Dutttagupta R, Willingham AT, et al. RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science*. 2007;316(5830):1484–8.
18. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, et al. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res*. 2012;22(9):1775–89.
19. Zhao X, Li J, Lian B, Gu H, Li Y, Qi Y. Global identification of *Arabidopsis* lncRNAs reveals the regulation of *MAF4* by a natural antisense RNA. *Nat Commun*. 2018;9(1):5056.
20. Xu W, Yang T, Wang B, Han B, Zhou H, Wang Y, et al. Differential expression networks and inheritance patterns of long non-coding RNAs in castor bean seeds. *Plant J*. 2018;95(2):324–40.
21. Qiu C-W, Zhao J, Chen Q, Wu F. Genome-wide characterization of drought stress responsive long non-coding RNAs in Tibetan wild barley. *Environ Exp Bot*. 2019;164:124–34.
22. Cagirici HB, Alptekin B, Budak H. RNA sequencing and co-expressed long non-coding RNA in modern and wild wheats. *Sci Rep*. 2017;7(1):10670.
23. Wang M, Yuan D, Tu L, Gao W, He Y, Hu H, et al. Long noncoding RNAs and their proposed functions in fibre development of cotton (*Gossypium* spp.). *New Phytol*. 2015;207(4):1181–97.
24. Zhao T, Tao X, Feng S, Wang L, Hong H, Ma W, et al. LncRNAs in polyploid cotton interspecific hybrids are derived from transposon neofunctionalization. *Genome Biol*. 2018;19(1):195.
25. Hu H, Wang M, Ding Y, Zhu S, Zhao G, Tu L, et al. Transcriptomic reporters depict the initiation of lint and fuzz fibres in cotton (*Gossypium hirsutum* L.). *Plant Biotechnol J*. 2018;16(5):1002–12.
26. Zou C, Wang Q, Lu C, Yang W, Zhang Y, Cheng H, et al. Transcriptome analysis reveals long noncoding RNAs involved in fiber development in cotton (*Gossypium arboreum*). *Sci China Life Sci*. 2016;59(2):164–71.
27. Lu X, Chen X, Mu M, Wang J, Wang X, Wang D, et al. Genome-wide analysis of long noncoding RNAs and their responses to drought stress in cotton (*Gossypium hirsutum* L.). *PLoS One*. 2016;11(6):e0156723.
28. Deng F, Zhang X, Wang W, Yuan R, Shen F. Identification of *Gossypium hirsutum* long non-coding RNAs (lncRNAs) under salt stress. *BMC Plant Biol*. 2018;18(1):23.
29. Wang G, Wang X, Zhang Y, Yang J, Li Z, Wu L, et al. Dynamic characteristics and functional analysis provide new insights into long non-coding RNA responsive to *Verticillium dahliae* infection in *Gossypium hirsutum*. *BMC Plant Biol*. 2021;21(1):68.
30. Zhang L, Wang M, Li N, Wang H, Qiu P, Pei L, et al. Long noncoding RNAs involve in resistance to *Verticillium dahliae*, a fungal disease in cotton. *Plant Biotechnol J*. 2018;16(6):1172–85.
31. Zhang J, Yang Z, Feng P, Zhong X, Ma Q, Su Q, et al. Identification and the potential roles of long non-coding RNAs in cotton leaves damaged by *Aphis gossypii*. *Plant Growth Regul*. 2019;88(3):215–25.
32. Lucero L, Ferrero L, Fonouni-Farde C, Ariel F. Functional classification of plant long noncoding RNAs: a transcript is known by the company it keeps. *New Phytol*. 2021;229(3):1251–60.
33. Budak H, Kaya SB, Cagirici HB. Long non-coding RNA in plants in the era of reference sequences. *Front Plant Sci*. 2020;11:276.
34. Bardou F, Ariel F, Simpson C, Romero-Barríos N, Laporte P, Balzergue S, et al. Long noncoding RNA modulates alternative splicing regulators in *Arabidopsis*. *Dev Cell*. 2014;30(2):166–76.
35. Franco-Zorrilla JM, Valli A, Todesco M, Mateos I, Puga MI, Rubio-Somoza I, et al. Target mimicry provides a new mechanism for regulation of microRNA activity. *Nat Genet*. 2007;39(8):1033–7.
36. Bari R, Pant BD, Stitt M, Scheible W-R. PHO2, microRNA399, and PHR1 define a phosphate-signaling pathway in plants. *Plant Physiol*. 2006;141(3):988–99.
37. Tang Y, Qu Z, Lei J, He R, Adelson DL, Zhu Y, et al. The long noncoding RNA FRILAIR regulates strawberry fruit ripening by functioning as a noncanonical target mimic. *PLoS Genet*. 2021;17(3):e1009461.
38. Ariel F, Jegu T, Latrasse D, Romero-Barríos N, Christ A, Benhamed M, et al. Noncoding transcription by alternative RNA polymerases dynamically regulates an auxin-driven chromatin loop. *Mol Cell*. 2014;55(3):383–96.
39. Dabbert T, Gore MA. Challenges and perspectives on improving heat and drought stress resilience in cotton. *J Cotton Sci*. 2014;18(3):393–409.
40. Loka DA, Oosterhuis DM, Ritchie GL. Water deficit stress in cotton. In: Oosterhuis DM, editor. *Stress physiology in cotton*. Cordova: The Cotton Foundation; 2011. p. 37–72.
41. Abdelraheem A, Esmaili N, O'Connell M, Zhang J. Progress and perspective on drought and salt stress tolerance in cotton. *Ind Crop Prod*. 2019;130:118–29.
42. Xu Z, Zhou G, Shimizu H. Are plant growth and photosynthesis limited by pre-drought following rewatering in grass? *J Exp Bot*. 2009;60(13):3737–49.
43. Zhang TZ, Hu Y, Jiang WK, Fang L, Guan XY, Chen JD, et al. Sequencing of allotetraploid cotton (*Gossypium hirsutum* L. acc. TM-1) provides a resource for fiber improvement. *Nat Biotechnol*. 2015;33(5):531–7.
44. Kim D, Perteau G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol*. 2013;14(4):R36.
45. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol*. 2010;28(5):511.
46. Kong L, Zhang Y, Ye ZQ, Liu XQ, Zhao SQ, Wei L, et al. CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Res*. 2007;35:W345–9.
47. Punta M, Coghill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, et al. The Pfam protein families database. *Nucleic Acids Res*. 2012;40:D290–301.
48. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta CT}$  method. *Methods*. 2001;25(4):402–8.
49. Yi X, Zhang Z, Ling Y, Xu W, Su Z. PNRD: a plant non-coding RNA database. *Nucleic Acids Res*. 2014;43:D982–9.
50. Dai X, Zhuang Z, Zhao PX. psRNATarget: a plant small RNA target analysis server (2017 release). *Nucleic Acids Res*. 2018;46:W49–54.
51. Chen Z, Jiang Q, Jiang P, Zhang W, Huang J, Liu C, et al. Novel low-nitrogen stress-responsive long non-coding RNAs (lncRNA) in barley landrace B968 (Liuzhoutouzidamai) at seedling stage. *BMC Plant Biol*. 2020;20(1):142.
52. Abdelraheem A, Adams N, Zhang J. Effects of drought on agronomic and fiber quality in an introgressed backcross inbred line population of upland cotton under field conditions. *Field Crops Res*. 2020;254:107850.
53. Kijun A, El-Dessouky HM, Benians T, Benians TAS, Goubet F, Meulewaeter F, et al. Analysis of the physical properties of developing cotton fibres. *Eur Polym J*. 2014;51:57–68.
54. Nadal M, Brodribb TJ, Fernández-Marín B, García-Plazaola JL, Arzac MI, López-Pozo M, et al. Differences in biochemical, gas exchange and hydraulic response to water stress in desiccation tolerant and sensitive fronds of the fern *Anemia cafferorum*. *New Phytol*. 2021;231(4):1415–30.
55. Snowden C, Ritchie G, Cave J, Keeling W, Rajan N. Multiple irrigation levels affect boll distribution, yield, and fiber micronaire in cotton. *Agron J*. 2013;105(6):1536–44.
56. Hou X, Fan J, Hu W, Zhang F, Yan F, Xiao C, et al. Optimal irrigation amount and nitrogen rate improved seed cotton yield while maintaining fiber quality of drip-fertilized cotton in northwest China. *Ind Crop Prod*. 2021;170:113710.
57. Gupta A, Rico-Medina A, Caño-Delgado AI. The physiology of plant responses to drought. *Science*. 2020;368(6488):266–9.
58. Ullah A, Sun H, Yang X, Zhang X. Drought coping strategies in cotton: increased crop per drop. *Plant Biotechnol J*. 2017;15(3):271–84.

59. Zhang M, Zheng XL, Song SQ, Zeng QW, Hou L, Li DM, et al. Spatiotemporal manipulation of auxin biosynthesis in cotton ovule epidermal cells enhances fiber yield and quality. *Nat Biotechnol.* 2011;29(5):453–8.
60. Zhao Q. Lignification: flexibility, biosynthesis and regulation. *Trends Plant Sci.* 2016;21(8):713–21.
61. Vogt T. Phenylpropanoid biosynthesis. *Mol Plant.* 2010;3(1):2–20.
62. Han LB, Li YB, Wang HY, Wu XM, Li CL, Luo M, et al. The dual functions of *WLM1a* in cell elongation and secondary wall formation in developing cotton fibers. *Plant Cell.* 2013;25(11):4421–38.
63. Gao Z, Sun W, Wang J, Zhao C, Zuo K. GhbHLH18 negatively regulates fiber strength and length by enhancing lignin biosynthesis in cotton fibers. *Plant Sci.* 2019;286:7–16.
64. Huang J, Chen F, Wu S, Li J, Xu W. Cotton *GhMYB7* is predominantly expressed in developing fibers and regulates secondary cell wall biosynthesis in transgenic *Arabidopsis*. *Sci China Life Sci.* 2016;59(2):194–205.
65. Duan J, Zhang M, Zhang H, Xiong H, Liu P, Ali J, et al. OsMIOX, a myo-inositol oxygenase gene, improves drought tolerance through scavenging of reactive oxygen species in rice (*Oryza sativa* L.). *Plant Sci.* 2012;196(196):143–51.
66. Li Z, Fu J, Shi D, Peng Y. Myo-inositol enhances drought tolerance in creeping bentgrass through alteration of osmotic adjustment, photosynthesis, and antioxidant defense. *Crop Sci.* 2020;60(4):2149–58.
67. Bashline L, Lei L, Li S, Gu Y. Cell wall, cytoskeleton, and cell expansion in higher plants. *Mol Plant.* 2014;7(4):586–600.
68. Liu S, Fukumoto T, Gena P, Feng P, Sun Q, Li Q, et al. Ectopic expression of a rice plasma membrane intrinsic protein (OsPIP1;3) promotes plant growth and water uptake. *Plant J.* 2020;102(4):779–96.
69. Li DD, Ruan XM, Zhang J, Wu YJ, Wang XL, Li XB. Cotton plasma membrane intrinsic protein 2s (PIP2s) selectively interact to regulate their water channel activities and are required for fibre development. *New Phytol.* 2013;199(3):695–707.
70. Wierzbicki AT, Blevins T, Swiezewski S. Long noncoding RNAs in plants. *Annu Rev Plant Biol.* 2021;72:245–71.
71. Tao X, Li M, Zhao T, Feng S, Zhang H, Wang L, et al. Neofunctionalization of a polyploidization-activated cotton long intergenic non-coding RNA DAN1 during drought stress regulation. *Plant Physiol.* 2021;186(4):2152–68.
72. Zhang X, Shen J, Xu Q, Dong J, Song L, Wang W, et al. Long noncoding RNA lncRNA354 functions as a competing endogenous RNA of miR160b to regulate ARF genes in response to salt stress in upland cotton. *Plant Cell Environ.* 2021;44(10):3302–21.
73. Khaksefidi RE, Mirlohi S, Khalaji F, Fakhari Z, Shiran B, Fallahi H, et al. Differential expression of seven conserved microRNAs in response to abiotic stress and their regulatory network in *Helianthus annuus*. *Front Plant Sci.* 2015;6:741.
74. Esmaeili F, Shiran B, Fallahi H, Mirakhorli N, Budak H, Martínez-Gómez P. In silico search and biological validation of microRNAs related to drought response in peach and almond. *Funct Integr Genomics.* 2017;17(2–3):189–201.
75. Ghorecha V, Zheng Y, Liu L, Sunkar R, Krishnaya NSR. MicroRNA dynamics in a wild and cultivated species of *Convolvulaceae* exposed to drought stress. *Physiol Mol Biol Plants.* 2017;23(2):291–300.
76. Fan G, Cao X, Niu S, Deng M, Zhao Z, Dong Y. Transcriptome, microRNA, and degradome analyses of the gene expression of *Paulownia* with phytoplasma. *BMC Genomics.* 2015;16:896.
77. Tai F-J, Wang X-L, Xu W-L, Li X-B. Characterization and expression analysis of two cotton genes encoding putative UDP-Glycosyltransferases. *Mol Biol.* 2008;42(1):44–51.
78. Hong YH, Meng J, He XL, Zhang YY, Luan YS. Overexpression of miR482c in tomato induces enhanced susceptibility to late blight. *Cells.* 2019;8(8):822.
79. Hong Y, Meng J, He X, Zhang Y, Liu Y, Zhang C, et al. Editing miR482b and miR482c simultaneously by CRISPR/Cas9 enhanced tomato resistance to *Phytophthora infestans*. *Phytopathology.* 2021;111(6):1008–16.
80. Zhu QH, Fan L, Liu Y, Xu H, Llewellyn D, Wilson I. MiR482 regulation of NBS-LRR defense genes during fungal pathogen infection in cotton. *PLoS One.* 2013;8(12):e84390.
81. Zhao YP, Shen JL, Li WJ, Wu N, Chen C, Hou YX. Evolutionary and characteristic analysis of RING-DUF1117 E3 ubiquitin ligase genes in *Gossypium* discerning the role of GhRDUF4D in *Verticillium dahliae* resistance. *Biomolecules.* 2021;11(8):1145.
82. Canto-Pastor A, Santos BAMC, Valli AA, Summers W, Schornack S, Baulcombe DC. Enhanced resistance to bacterial and oomycete pathogens by short tandem target mimic RNAs in tomato. *Proc Natl Acad Sci U S A.* 2019;116(7):2755–60.
83. Tuteja N, Tarique M, Banu MSA, Ahmad M, Tuteja R. *Pisum sativum* p68 DEAD-box protein is ATP-dependent RNA helicase and unique bipolar DNA helicase. *Plant Mol Biol.* 2014;85(6):639–51.
84. Wang K, Chen H, Ortega-Perez M, Miao Y, Ma Y, Henschen A, et al. Independent parental contributions initiate zygote polarization in *Arabidopsis thaliana*. *Curr Biol.* 2021;50960-9822(21):01137–4.
85. Li H, Yang Y, Wang H, Liu S, Jia F, Su Y, et al. The receptor-like kinase *ERECTA* confers improved water use efficiency and drought tolerance to poplar via modulating stomatal density. *Int J Mol Sci.* 2021;22(14):7245.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

