Genome-wide analysis of PRR gene family uncovers their roles in circadian rhythmic changes and response to drought stress in *Gossypium hirsutum* L.

Jingjing Wang [Equal first author, 1, 2], **Zhaohai Du** [Equal first author, 1], **Xuehan Huo** [1, 2], **Juan Zhou** [1], **Yu Chen** [1], **Jingxia Zhang** [1], **Ao Pan** [1], **Xiaoyang Wang** [3], **Furong Wang** [Corresp., 1, 2], **Jun Zhang** [Corresp. 1, 2]

[1] Key Laboratory of Cotton Breeding and Cultivation in Huang-Huai-Hai Plain, Ministry of Agriculture and Rural Affairs, Cotton Research Center, Shandong Academy of Agricultural Sciences, Jinan, P. R. China

[2] College of Life Sciences, Shandong Normal University, Jinan, P. R. China

[3] State Key Laboratory of Cotton Biology State Key Laboratory of Cotton Biology, Institute of Cotton Research, Chinese Academy of Agricultural Sciences, Anyang, P. R. China

Corresponding Authors: Furong Wang, Jun Zhang
Email address: wfr1125@126.com, zj0928@126.com

Background. The circadian clock not only participates in regulating various stages of plant growth, development and metabolism, but confers plant environmental adaptability to stress such as drought. Pseudo-Response Regulators (PRRs) are important component of the central oscillator (the core of circadian clock) and play a significant role in plant photoperiod pathway. However, no systematical study about this gene family has been performed in cotton.

Methods. *PRR* genes were identified in diploid and tetraploid cotton using bioinformatics methods to investigate their homology, duplication and evolution relationship. Differential gene expression, KEGG enrichment analysis and qRT-PCR were conducted to analyze *PRR* gene expression patterns under diurnal change and their response to drought stress.

Results. A total of 44 PRR family members were identified in four *Gossypium* species, with 16 in *G. hirsutum*, 9 in *G. barbadense* as well as in *G. arboreum* and 10 in *G. raimondii*. Phylogenetic analysis indicated that PRR proteins were divided into five subfamilies and whole genome duplication or segmental duplication contributed to the expansion of *Gossypium* PRR gene family. Gene structure analysis revealed that members in the same clade are similar, and multiple cis-elements related to light and drought stress response were enriched in the promoters of *GhPRR* genes. qRT-PCR results showed that *GhPRR* genes transcripts present four expression peaks (6 h, 9 h, 12 h, 15 h) during 24 hours and form obvious circadian waves. Transcriptome data with PEG treatment, along with qRT-PCR verification suggested that members of clade III (*GhPRR5a-d*) and clade V (*GhPRR3a* and *GhPRR3c*) may be involved in drought response. This study provides an insight into understanding the function of *PRR* genes in circadian rhythm and in response to drought stress in cotton.

Genome-wide analysis of PRR gene family uncovers their roles in circadian rhythmic changes and response to drought stress in *Gossypium hirsutum* L.

Jingjing Wang[1,2], Zhaohai Du[1], Xuehan Huo[1,2], Juan Zhou[1],Yu Chen[1], Jingxia Zhang[1], Ao Pan[1], Xiaoyang Wang[3], Furong Wang[1,2*], Jun Zhang[1,2*]

1. Key Laboratory of Cotton Breeding and Cultivation in Huang-Huai-Hai Plain, Ministry of Agriculture and Rural Affairs, Cotton Research Center, Shandong Academy of Agricultural Sciences, Jinan, P. R. China

2. College of Life Sciences, Shandong Normal University, Jinan, P. R. China

3. State Key Laboratory of Cotton Biology, Institute of Cotton Research, Chinese Academy of Agricultural Sciences, Anyang, P. R. China

*Corresponding Author:

1. Furong Wang:

[1]Gongyebei Road, Jinan, Shandong Province, 250100, China.

[2] Wenhuadong Road, Jinan, Shandong Province, 250014, China.

E-mail: wfr1125@126.com; wangfurong@shandong.cn

2. Jun Zhang:

[1]Gongyebei Road, Jinan, Shandong Province, 250100, China.

[2] Wenhuadong Road, Jinan, Shandong Province, 250014, China.

E-mail: zj0928@126.com; mhzxzhangjun@shandong.cn

Abstract

Background. The circadian clock not only participates in regulating various stages of plant growth, development and metabolism, but confers plant environmental adaptability to stress such as drought. Pseudo-Response Regulators (PRRs) are important component of the central oscillator (the core of circadian clock) and play a significant role in plant photoperiod pathway. However, no systematical study about this gene family has been performed in cotton.

Methods. *PRR* genes were identified in diploid and tetraploid cotton using bioinformatics methods to

investigate their homology, duplication and evolution relationship. Differential gene expression, KEGG enrichment analysis and qRT-PCR were conducted to analyze *PRR* gene expression patterns under diurnal change and their response to drought stress.

Results. A total of 44 PRR family members were identified in four *Gossypium* species, with 16 in *G. hirsutum*, 9 in *G. barbadense* as well as in *G. arboreum* and 10 in *G. raimondii*. Phylogenetic analysis indicated that PRR proteins were divided into five subfamilies and whole genome duplication or segmental duplication contributed to the expansion of *Gossypium* PRR gene family. Gene structure analysis revealed that members in the same clade are similar, and multiple cis-elements related to light and drought stress response were enriched in the promoters of *GhPRR* genes. qRT-PCR results showed that *GhPRR* genes transcripts present four expression peaks (6 h, 9 h, 12 h, 15 h) during 24 hours and form obvious circadian waves. Transcriptome data with PEG treatment, along with qRT-PCR verification suggested that members of clade III (*GhPRR5a-d*) and clade V (*GhPRR3a* and *GhPRR3c*) may be involved in drought response. This study provides an insight into understanding the function of *PRR* genes in circadian rhythm and in response to drought stress in cotton.

Keywords: *Gossypium hirsutum*; PRR family; photoperiod; circadian rhythm; drought response

Introduction

The circadian clock is an autonomous endogenous biological rhythm that enables the living organisms to adapt to external daily and seasonal cycles, which play a significant role in plant growth and development for plant fitness (*Harmer, 2009*; *Hsu et al., 2014*; *Lee et al., 2005*; *McClung, 2006*; *Uehara et al.,* 2019) Although the circadian clock in different organisms is tissue-specific, most organisms have a conserved molecular mechanism - the core oscillator of positive and negative feedback loops formed at both the transcriptional and translational levels based on genome-wide gene expression regulation. (*Harmer, 2009*; *Hsu PY et al., 2014*; *Takata et al., 2009*; *Uehara et al., 2019*).Numerous studies have indicated that imperative roles for PRR gene family (PRR9, PRR7, PRR5, PRR3 and TOC1) in circadian clock (*Eriksson et al., 2003*; *Farre et al., 2005*; *Gould et al., 2006*; *Ito et al., 2009*; *Kaczorowski et al., 2003*; *Nakamichi et al., 2020*; *Salome et al., 2005*; *Yamamoto et al., 2003*).

In *Arabidopsis thaliana*, the gene expression and protein expression levels of PRR family members have obvious circadian rhythmic expression pattern (*Matsushika et al., 2000*). PRR proteins contain two domains,

the N-terminal contains a conserved REC (Response regulator receiver) domain, the C-terminus is a CCT domain, and CCT domain might interacts with CONSTITUITIVE PHOTOMOR-PHOGENIC 1 (COP1) to control CONSTANS (CO) protein stability (Jang *et al*., 2008). PRRs could interact with CO at specific times and stabilize CO expression during the day, which promoting the CO protein to bind the promoter of *FLOWERING LOCUS T* (*FT*), inducing FT expression and promoting flowering (*Hayama et al., 2017*; *Kobayashi et al., 1999*). Many studies showed that *PRRs* have circadian rhythmic expression levels in both transcriptional and protein levels, whether in continuous light or dark (*Más et al., 2003*; *Strayer et al., 2000*). Either in the *toc1* deletion mutant or *TOC1* overexpressing plants of *Arabidopsis thaliana*, the performance of the core oscillator has significant changes (*Huang et al., 2012*). Besides, PRR9, PRR7 and PRR5 could act as transcriptional repressors of CCA1 and LHY (*Nakamichi et al., 2010*).

At present, research mainly focuses on exploring the molecular mechanism of the photoperiod regulation pathway in *Arabidopsis thaliana*, and its regulation mechanism is becoming clear (*Song et al., 2013*; *Wang et al., 2013*; *Wickland et al., 2015*). Flowering time is an important factor affecting crop yield, thus dissection of photoperiod pathways regulating flowering time in crops and ornamental plants also becomes one of the hotspots in current researches (*Brambilla et al., 2017*; *Nakamichi et al., 2015*; *Yang et al., 2020*). However, molecular details concerning the photoperiodic control of crop flowering remain unclear. Only some studies on the cloning and function analysis of PRR genes have been carried out in crops currently, such as rice (*Oryza sativa*) (*Murakami et al., 2005*), wheat (*Triticum aestivum*) (*Nakahira et al., 1998*), barley (*Hordeum vulgare*) (*Turner et al., 2005*) and soybean (*Glycine max*) (*Liu et al., 2009*).

Flowering at the appropriate period has an essential effect on the fiber yield and quality of cotton, and there were only a few studies on genes related to flowering regulation in cotton (*Gossypium spp*.) (*Cai et al., 2017*; *Zhang et al., 2016*). With the completion of the genome sequencing of *Gossypium* species (*Hu et al., 2019*; *Huang et al., 2020*; *Li et al., 2015*; *Wang et al., 2012*; *Wang et al., 2018*; *Yuan et al., 2015*; *Zhang et al., 2015*), the identification of new genes and the establishment of a new regulatory model would be helpful for studying the function of genes involved in cotton flowering pathway.

In addition, the biological clock plays a vital role in adapting to external environmental stress, such as drought stress. In soybeans, studies shown that drought stress affects the expression of circadian clock genes,

and the expression of drought-responsive genes also has circadian rhythm (*Gomes et al., 2014*). TOC1 has been shown to directly bind to the ABAR promoter region and regulate the periodic expression of ABAR, while ABA can up regulate TOC1.Therefore, TOC1 is considered to act as a molecular switch between the drought stress signaling pathway and the biological clock (*Legnaioli et al., 2009*).

Here, we identified 44 *PRR* genes from the four *Gossypium* species, and conducted basic bioinformatics analysis. We also investigated the periodic expression pattern of *PRR* family members at the transcriptional level during 24 hours. Further, we identified six *PRR* members respond to drought stress by analyzing transcriptome data with PEG treatment along with qRT-PCR verification. This study lays a foundation for studying the molecular mechanism of cotton photoperiod regulation and also provides an insight into understanding *PPRs* gene function in response to drought stress in cotton.

Materials & methods

Identification of PRR gene family in *Gossypium spp.*

The domain numbered PF00072 (Response regulator receiver domain) and PF06203 (CCT motif) in the Pfam database (*EI-Gebali et al., 2019*) are the domain of plant light signal transcription factor. Firstly, the genome-wide information files of *G. hirsutum*, *G. barbadense*, *G. arboreum* and *G. raimondii* are downloaded from the cotton genome database (http://mascotton.njau.edu.cn/info/1054/1118.htm).Then, candidate sequences REC and CCT domains were aligned and searched in HMMER 3.0 (https://www.ebi.ac.uk/Tools/hmmer/) (*Jacob et al., 2007*) and the BLAST program. Next, the sequences obtained were searched and verified on the Conserved Domain database CDD (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) and SMART (http://smart.embl-heidelberg.de/) (*Letunic et al., 2002*), to determine its conserved domain.

Finally, the whole genome sequence of *Arabidopsis thaliana* (*Initiative et al., 2000*), Cocoa (*Theobroma cacao*) (*Argout et al., 2011*) and rice (*Oryza sativa*) (*Yu et al., 2005*) were loaded from the Arabidopsis database TAIR10 (https://www.arabidopsis.org/), the plant genome database Phytozome 12 (http://phytozome.jgi.doe.gov/pz/portal.html) and EnsemblPlants (http://plants.ensembl.org/index.html), respectively. Finally, the online site ExPASy Proteomics Server (http://www.expasy.org/) and Softberry(http://linux1.softberry.com/berry.phtml?topic=protcomppl &group=programs&subgroup=proloc) were used to analyze the physicochemical properties of the identified cotton PRR gene family, including

amino acid number, nucleotide data, molecular weight, isoelectric point prediction and subcellular localization.

Chromosomal locations, duplications, and synteny analysis of PRR gene members

Chromosomal location information for *PRR* genes was obtained from a gff3-file of each cotton genomic databases and genes were mapped on the chromosomes using TBtools (https://github.com/CJ-Chen/TBtools). Then MCScanX (*Wang et al., 2012*) was used to determine and analyze cotton *PRR* duplication and collinearity, Circos (http://circos.ca/) and Adobe illustrator CS5 software were used to conducted image showing gene location and gene homology relationship.

Phylogenetic analyses and gene structure organization of the PRR proteins in *Gossypium spp.*

To analyze evolutionary relationship, multi-protein sequence alignment of the PRR proteins were aligned using MEGA7 (*Sudhir et al., 2016*), and constructed a phylogenetic tree with neighbor-joining (NJ) method, with the bootstrap 1000. Finally, the evolutionary tree is visualized and beautified by the online software iTOL (https://itol.embl.de/) (*Letunic et al., 2019*).

SeqHunter1.0 (*Ye et al., 2010*) and TBtools software were used to extract *PRR*s gene families GFF3 files from upland cotton genome sequencing file. Location information and the predicted CDS sequence of the *PRR* members were obtained**,** and presenting the gene structures in the online software Gene Structure Display Server (GSDS 2.0) (http://gsds.cbi.pku.edu.cn/index.php) (*Guo et al., 2007*), Then we performed motifs analysis by the online software MEME (http://meme-suite.org/) (*Bailey et al., 2009*) with following parameters: the maximum number discovered for the motif is set 1 to 10, and the other parameters are default values. The graphic display is based on the Amazing optional gene viewer section in the software TBtools.

To identify the cis-elements in the promoter sequences of the 16 PRR family genes in *G. hirsutum*, the 2000 bp of genomic sequences upstream of the start codon of each *PRR* gene were submitted to the online site PlantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) to predict the cis-element type, and the results are displayed by the Simple Bio Sequence Viewer in TBtools.

Plant materials and treatment

The upland cotton (*G. hirsutum*) accession (Lumianyan 19, LMY 19) selected in this study were kept in our laboratory, planted in growth chamber (day/night temperature cycle of 28°C light/25 °C dark with a 12-photoperiod), and samples were picked every 3h from leaf in three-true-leaves stage. Germinated TM-

1 cotton seeds were planted in the same photoperiod and temperature environment as LMY19, and treated with 400 mM PEG at the three-leaf stage, samples collected at 0, 1, 3, 6 h respectively. Three biological replicates of each sample were performed and all samples were freezed with liquid nitrogen immediately and stored at -80 °C for qRT-PCR.

RNA isolation and qRT-PCR analysis

The RNA was extracted from the samples using the Rapid Universal Plant RNA Extraction Kit (Huayueyang Biotechnology Co. Ltd.), and the Prime Scrip First Strand cDNA Synthesis Kit (Takara) used for reverse transcription, SYBR Premix Ex Taq II. (Takara) kit used for real-time PCR experiment, qRT-PCR analysis was carried out using SYBR Green on the Roche LightCycler® 480 II. The primers of PRR gene family were designed using Primer Premier 5.0 software and listed in table S1, and the actin gene (AF059484) was selected as the internal reference gene (*Zhang et al., 2013*). The volume of the qRT-PCR reaction was 20 μL, and the amplification procedure was as follows: pre-denaturation at 95 ℃ for 30 s; denaturation at 95 ℃ for 5 s, annealing at 60℃ for 30 s, 40 cycles. Three biological and technical replicates were performed for the qRT-PCR tests. The relative gene expression levels were quantified by the $2^{-\Delta\Delta Ct}$ method (*Livak et al., 2001*).

Expression patterns and pathway enrichment analysis of PRR members

The transcriptome data sets of *G. hirsutum* TM-1 were downloaded from SRA database (PRJNA490626) in NCBI (https://www.ncbi.nlm.nih.gov/sra/) (*Zhang et al., 2015*), and the accession numbers are SRS824360, SRS824361, and SRS824362 respectively. The gene expression pattern of *PRR* genes were displayed by R/heatmap with the expression values normalized by $\log_2$(FPKM+1) and the differentially expressed genes (DEGs) were identified using the R/mpoly (*Kahle et al., 2013*). KEGG analyses of DEGs were conducted in the Kyoto Encyclopedia of Genes and Genomes (KEGG) database for enrichment (*Kanehisa et al., 2014*), KEGG enrichment of DEGs was evaluated with KOBAS2.0 software (*Xie et al., 2011*) and bubble graph was displayed by R/ggplot2.

Results

Genome-wide identification of PRR family genes in *Gossypium spp.*

Based on multiple sequence alignment analysis, complete *PRR* genes were identified in four *Gossypium* species, including 16 in *G. hirsutum* (AD_1), 9 in *G. arboretum* (A_2), 10 in *G. raimondii* (D_5), and 9 in *G.*

barbadense (AD_2). Additionally, we proceeded with *PRR* genes retrieved from plant genome database, with 5 in *Arabidopsis* (dicots), 5 in *rice* (monocots), and 6 in *cocoa* (dicot). All of them were renamed based on the homologous genes in Arabidopsis (Table S2). The number of *PRR* gene family in *G. hirsutum* was about twice as that in *G. arboreum* (A group) or *G. raimondii* (D group), consistent with they are tetraploid and diploid genomes respectively. The basic information of *PRR* genes including protein sequence length, isoelectric points, and molecular weight in cotton were listed in Table S3. The predicted GhPRR proteins ranged from 553 (GhPRR1b) to 796 (GhPRR3a) amino acids, with isoelectric points changed from 5.16 (GhPRR9a) to 8.54 (GhPRR7c) and molecular weight from 61.87 kDa (GhPRR1a) to 85.93 kDa (GhPRR3a).

Chromosomal locations, duplications, and synteny analysis of PRR gene members

In order to display the chromosome distribution of *PRR* genes, mapping them on the corresponding chromosome. Eight of *GhPRR* genes were located on chromosomes of At sub-genome while five of *GhPRR* genes were on that of Dt sub-genome and three *GhPRR* genes were present in different scaffolds (Fig.S1). We further conducted whole genome collinearity analysis of 44 identified *PRR* genes in cotton, and explored the locus relationships between At and Dt sub-genomes as well as with A and D diploid cotton genomes (Fig.1A, Table S4). There are 34 orthologous gene pairs were resulted from whole genome duplication or segmental duplication among *Gossypium spp.*.Whole Genome duplication or segmental duplication was suggested to be the main causes of PRR gene family expansion in cotton (Table S5).

Phylogenetic analyses and gene structure organization of the PRR proteins in *Gossypium spp.*

To investigate the evolutionary relationship of GhPRR proteins among mentioned seven species, phylogenetic tree was constructed (Fig.1B). The PRR family of *Gossypium* was divided into 5 subgroups (clade I-V). There were 13 PRRs in Clade III (three GaPRRs, GbPRRs and GrPRRs respectively, four GhPRRs) and 11 PRRs (one GrPRR, two GaPRRs, four GbPRRs and GhPRRs individually) in clade IV. Clade I consisted of 9 PRRs (one GaPRR, two GbPRRs and GhPRRs singly, four GrPRRs), Clade V contained 7 PRRs (one GrPRR, two GaPRRs and four GhPRRs) and Clade II had 4 PRRs (one GaPRR and GrPRR respectively, two GhPRRs). GhPRRs were distributed throughout five subgroups (clade I-V), clade-I, clade-II and clade-IV containing PRRs from monocots and dicots simultaneously, illustrating that evolution of *GhPRR* genes in three clades occurred before the separation of monocots and dicots.

PRRs protein in *G. hirsutum* was also divided into five subgroups (Fig. 2A), consistent with phylogenetic analyses. The motif distribution indicated that the order, size, and location of the motifs in the same subgroup were similar, but there were significant variety between different subgroups. Among them, 37.5% of the family members have the same sequence of motif structure: motif 4_9_3_1_7_5_6_10_8_2, while Clade-I contains the least number of motifs with only 5 motifs. All members of the PRR gene family contain motif1, motif2, motif3, motif4 and motif6, which are the conserved motifs of PRR family. In addition, the gene structure analysis exhibited that the distribution of introns and exons were similar among different subgroups, and the functional elements REC and CCT were distributed in both end side of each gene. All of member contained three REC structure elements, and most member contain two CCT domains, except that two members of the Clade-I subgroup contain one CCT domain.

To further analyze the transcriptional regulation and potential function of the *PRR* genes, the cis-elements in the promoter region were predicted (Fig. 2B).The results displayed that there are abundant regulatory elements existing in the promoter region, mainly focused on light response elements (G-Box, GT1-motif and TCT-motif, etc.), hormone responsive elements: abscisic acid response (ABRE), MeJA-response (CGTCA-motif and TGACG-motif), gibberellin-responsive element (TATC-box, P-box and GARE-motif), and stress responsive elements: drought-inducibility (MBS), low-temperature response (LTR), etc. These results implied that the expression of *GhPRR* genes may be induced by light, abscisic acid and drought stress.

The periodic expression pattern of PRR members under diurnal change

A feature shared by many clock gene transcripts is that their abundance is subject to diurnal oscillation. To analyze the peak transcripts of *GhPRR*s under diurnal cycle, the relative expression levels of *GhPRRs* together with its related genes (*GhFT* (*Gh_A08G2015*), *GhCO* (*Gh_D12G0543*), *GhLHY* (*Gh_D11G1068*) and *GhCCA1* (*Gh_A05G2124*)) under circadian rhythm was detected by qRT-PCR (Fig. 3), the results presented that the expression of *GhPRR* genes reached the peak every 3 hours after 6 hours of light condition, and there were multiple members at each peak. *GhFT*, *GhCO*, and *GhLHY* had the peak expression at 3 hours of light. Subsequently, members include-II (*GhPRR9a* and *GhPRR9b*) and clade-IV (*GhPRR7a*, *GhPRR7b*, *GhPRR7c* and *GhPRR7d*) reached the expression peak after 6 hour of light condition, and then members in clade III (*GhPRR5a*, *GhPRR5b*, *GhPRR5c* and *GhPRR5d*) and clade-V (*GhPRR3a* and *GhPRR3c*) at 9 hour, another two members of clade-V (*GhPRR3b* and *GhPRR3d*) at 12 hour. Finally, members (*GhPRR1a* and *GhPRR1b*)

in clade-I reached expression peak at night. Additionally, the expression of *GhLHY* and *GhPRR1b* always showed an opposite trend during 24 hours, it can be speculated that a mutual inhibition maybe exist between the two genes. This results indicated that expression of *GhPRR* genes has obvious circadian waves during 24 hours.

Identification of drought-stress related PRR genes in *G. hirsutum*

To investigate the roles for PRR genes in response to drought, we investigated the expression profile of *GhPRRs* under polyethylene glycol (PEG) treatment at 1, 3 and 6 h from the published transcriptome data sets. We selected the group data at 6 h treated with PEG for KEGG (Kyoto Encyclopedia of Genes and Genomes) analysis for it having the largest number (848) of significantly up-regulated genes (Fig.4A).The results revealed that the up-regulated genes are mainly involved in circadian rhythm, photosynthesis, starch and sucrose metabolism, etc. (Fig.4B). Six of *GhPRR* genes including four members in clade III (*GhPRR5a-d*) and two in clade-V (*GhPRR3a* and *GhPRR3c*) were involved in circadian rhythm pathway. The expression pattern of *PRR* genes under PEG treatment showed that these six members have same expression trend and high expression level at 6 h with PEG treatment (Fig. 4C and Table S6), suggesting these *PRR* genes are significantly induced by PEG treatment.

To further prove the expression changes of these genes at different time of PEG treatment (1h, 3h, 6 h), the expression level of all member of PRR family were detected by qRT-PCR (Fig.5). All PRR genes displayed almost the similar expression changes compared with transcriptome data sets. The expression of the above mentioned genes (*GhPRR5a-d*, *GhPRR3a* and *GhPRR3c*) at the sixth hour after PEG treatment was significantly higher than that of the blank control. Especially, the expression level of *GhPRR3c* gene is almost nine times higher than that of the blank control, suggesting these genes may be involved in drought response.

Discussion

Light, one of the vital environmental factors, plays a significant role in promoting plant growth and development. Especially, with the alternating of sunrise and sunset, plants form a unique biological clock to regulate the growth and metabolic activities, like regulation of flowering time (*Hayama et al., 2017*; *Song et al., 2015*), hypocotyl elongation (*Seaton et al., 2015*; *Soy et al., 2016*; *Zhu et al., 2016*), biotic (*Bhardwaj et al., 2011*; *Korneli et al., 2014*; *Zhang et al., 2013*) and abiotic (*Keily et al., 2013*; *Nakamichi et al., 2009*) stress

response, and so on. Circadian clock related genes regulate many vital agronomic traits, especially in the photoperiod pathway of plants. There are a few studies about *PRR* genes in cotton (*Gossypium spp*.), however, advances in cotton genomics and genetics recent years allowed us to perform a systematic study on *PRR* genes and to probe their potential functions in circadian clock. This study will provide basic information and for further investigation of cotton *PRR* gene functions.

Here, 16 *GhPRR* genes were identified totally in *G. hirsutum*, which are homologous with 5 PRR genes in *Arabidopsis* (*PRR1* (*TOC1*), *PRR3*, *PRR5*, *PRR7*, *PRR9*). According to chromosomal localization and genomic collinearity analysis, it can be speculated that due to the hybridization of A and D subgenome in the *G. hirsutum*, the gene amplification is carried out by tandem repeat and fragment replication, resulting in more family members (*Jackson et al., 2010*). There is a high degree of collinearity between the genes of the A subgenome and the D subgenome of the tetraploid *G. hirsutum* (*Li et al., 2015*). In this study, 14 (7 pairs) of 16 PRR members are orthologous genes, indicating that *G. hirsutum* has undergone large-scale gene rearrangement at the genomic level during species formation, which is consistent with the results of the allotetraploid *G. hirsutum* genome (*Wang et al., 2018*; *Li et al., 2015*; *Zhang et al., 2015*).

A large number of experimental studies have been carried out about circadian clock in *Arabidopsis* (*Alabadí et al., 2001*; *Más et al., 2003*; *Legnaioli et al., 2009*). PRRs proteins interact with CCA1 and LHY through complex mechanisms, playing a vital role in the growth and development, flowering induction and metabolic regulation of plants (*Harmer, 2009*; *Legnaioli et al., 2009*; *Mizuno et al., 2005*). In addition to confirming the regulation of flowering by the circadian clock in model plants, the study of circadian clock mechanism has been carried out in crops recently. The function of some circadian clock-related genes has been cloned and verified based on gene homology in major crops, such as rice, soybean (*Gome et al., 2014*; *Xue et al., 2012*; *Yang et al., 2013*). In rice, overexpressing *OsPRR37* weakens the transcriptomic rhythms and alters the phases of rhythmic genes. So far, circadian clock regulation mechanism in cotton is still a mystery, only one study has identified *Gh_D03G0885* (*GhPRR1b*) as a candidate gene for cotton early maturity traits using genotyping-by-sequencing (*Li et al., 2017*). TOC1 (known as Pseudo Response Regulator, PRR1) is an important component of the core oscillator and closed positive and negative feedback loop with LHY (Late elongated hypocotyl) and CCA1 (circadian clock associated 1), formulating the basic framework of the *Arabidopsis* circadian clock core

oscillator (*Alabadí et al., 2001*; *Gendron et al., 2012*; *Huang et al., 2012*). As an inhibitor of circadian clock gene expression, TOC1 gene can inhibit the expression of most circadian clock core genes, and affect flowering pathway of photoperiod regulation by controlling the function of circadian clock (*Strayer et al., 2000*; *Pokhilko et al., 2012*). However, the function and regulatory pathways involved of this gene have not been reported in cotton.

Further, qRT-PCR analysis revealed that the relative expression of PRR members had apparent circadian waves among 24 hours, which similar with that of PRR members *(PRR1/TOC1, PRR3, PRR5, PRR7, PRR9)* in *Arabidopsis*. Transcript expression peaks appear in the order of *PRR9*, *PRR7*, *PRR5*, *PRR3* and *TOC1* (*PRR1*) in *Arabidopsis* (*Matsushika et al., 2000*), while four expression peaks appeared in this study and there were multiple members at each peak, speculating that it is related to chromosome doubling in the process of forming allotetraploid in *G. hirsutum* (*Jackson et al., 2010*). The *PRR1a* gene had the last peak of expression and highly expressed at night, which consistent with that of *APRR1* in *Arabidopsis* (*Caluwé et al., 2016*), while *GhPRR1b* has two peak of expression at night in this study. Therefore, detailed study should be carried out about this the gene in cotton.

In cotton, the expression peak of the members of the *GhPRR* gene family follows the expression peak of *GhLHY*, which is consistent with the results in *Arabidopsis* (*Mizuno et al., 2005*). *GhPRR1b* and *GhLHY* have opposite expression trends among 24 hours, and the *GhPRR1b* gene has high homology with *PRR1* in Arabidopsis thaliana by alignment, so it is speculated that *GhPRR1b* is the core component of the circadian clock in *G. hirsutum*, and there is maybe a mutual inhibition between *GhPRR1* and *GhLHY*. As an important factor in the export pathway of the circadian clock, CO protein has been proved in *Arabidopsis* to confirm the stability of PRRs protein-mediated CO expression, and can enhance the binding of CO to *FT* promoter, then *FT* start transcribe and promote flowering (*Jang et al., 2008*). The pathway of PRRs family members mediate the stability of CO expression still needs further experiments in cotton.

In addition, there are many studies focus on the response of circadian clock to abiotic stress in crops (*Flowers et al., 2004*; *Lu et al., 2017; Zhang et al., 2020*). TOC1 can bind to the *ABAR* promoter of ABA-related genes and regulate its circadian rhythm expression, and can be thought to act as molecular switches between drought stress signaling pathways and circadian clocks in *Arabidopsis* (*Legnaioli et al., 2009*). In

soybeans, studies have also shown that drought stress affects the expression of circadian clock genes, and the expression of drought-responsive genes also has circadian rhythm (*Gomes et al., 2014*). Based on these researches, this study identified 16 PRR members in cotton and analyzed the expression pattern of PRR genes among 24 hours and in response to drought stress, the result showed that PRR members expression display obvious circadian waves and six of them may be involved in responding to drought stress, which is helpful to understand the evolution and function of the PRRs gene family, and provide thoughts and clues for further study the function of the PRR gene family in cotton.

Conclusions

In this study, we identified 44 *PRR* genes in cotton (*Gossypium spp.*) and classified them into 5 subgroups based on the phylogenetic tree. Then comprehensively and systematically analyzed PRRs in cotton (*Gossypium spp*.), including the domains, the gene structure, promoter cis-acting element, chromosome localization distribution and collinearity analysis. In addition, we also investigated the evolutionary relationship of PRRs among *G. hirsutum, G. barbadense, G. arboreum* and *G. raimondii*, *Arabidopsis thaliana*, *Theobroma cacao* and *Oryza sativa*. Moreover, qRT-PCR results showed that the expression of members of PRRs family has obvious circadian waves, and Gene differential expression and KEGG enrichment analysis of the transcriptome data with PEG treatment, along with qRT-PCR verification altogether demonstrated members of clade III (*GhPRR5a-d*) and two members of clade-V (*GhPRR3a* and *GhPRR3c*) are significantly induced by PEG treatment, so it is speculated that these *GhPRR* genes may be involved in drought response. This study will provide a theoretical basis for studying the function of *PRRs* in cotton.

Availability of data and materials

All related data are available within the manuscript and its additional files. The RNA sequences raw data was downloaded from the SRA database, National Center for Biotechnology Information (NCBI) under the accession numbers (PRJNA490626).

Funding

This work is supported by the National Science Foundation in China (31671742); the National Project of Modern Agricultural Industry Technology System in China [CARS-15-05]; the Taishan Scholars Program of

Shandong Province [ts201511070].

Acknowledgements

The authors thank Guoyong Fu for kindly revising this manuscript.

Conflict of Interest

The authors declare that they have no conflict of interest.

Reference

Alabadí D, Oyama T, Yanovsky MJ, Harmon FG, Más P, Kay SA. 2001. Reciprocal regulation between TOC1 and LHY/CCA1 within the *Arabidopsis* circadian clock. *Science* **293**:880-883 DOI 10.1126/science.1061320.

Argout X, Salse J, Aury JM, Guiltinan MJ, Droc G, Gouzy J, Allegre M, Chaparro C, Legavre T, Maximova SN, Abrouk M, Murat F, Fouet O, Poulain J, Ruiz M, Roguet Y, Rodier-Goud M, Barbosa-Neto JF, Sabot F, Kudrna D, Ammiraju JS, Schuster SC, Carlson JE, Sallet E, Schiex T, Dievart A, Kramer M, Gelley L, Shi Z, Bérard A, Viot C, Boccara M, Risterucci AM, Guignon V, Sabau X, Axtell MJ, Ma Z, Zhang Y, Brown S, Bourge M, Golser W, Song X, Clement D, Rivallan R, Tahi M, Akaza JM, Pitollat B, Gramacho K, D'Hont A, Brunel D, Infante D, Kebe I, Costet P, Wing R, McCombie WR, Guiderdoni E, Quetier F, Panaud O, Wincker P, Bocs S, Lanaud C. 2011. The genome of *Theobroma cacao*. *Nature Genetics* **43**:101-108 DOI 10.1038/ng.736.

Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren JY, Li WW, Noble WS. 2009. MEME Suite: tools for motif discovery and searching. *Nucleic Acids Research* **37**:202 DOI 10.1093/nar/gkp335.

Bhardwaj V, Meier S, Petersen LN, Ingle RA, Roden LC. 2011. Defence responses of *Arabidopsis thaliana* to infection by *Pseudomonas syringae* are regulated by the circadian clock. *PLoS ONE* **6**: e26968 DOI 10.1371/journal.pone.0026968.

Brambilla V, Gomez-Ariza J, Cerise M, Fornara F. 2017. The importance of being on time: regula tory networks controlling photoperiodic flowering in cereals. *Frontiers in Plant Science* **8**:665 DOI 10.3389/fpls.2017.00665.

Cai D, Liu H, Sang N, Huang X. 2017. Identification and characterization of CONSTANS-like (COL) gene family in upland cotton (*Gossypium hirsutum* L.). *PLoS ONE* **12**:e0179038 DOI 10.1371/journal.pone.0179038.

De Caluwé J, Xiao Q, Hermans C, Verbruggen N, Leloup JC, Gonze D. 2016. A compact model for the complex plant circadian clock. *Frontiers in Plant Science* **7**:74 DOI 10.3389/fpls.2016.00074.

Du XM, Huang G, He SP, Yang ZE, Sun GF, Ma XF, Li N, Zhang XY, Sun JL, Liu M, Jia YH, Pan ZE, Gong WF, Liu ZH, Zhu HQ, Ma L, Liu FY, Yang DG, Wang F, Fan W, Gong Q, Peng Z, Wang LR, Wang XY, Xu SJ, Shang HH, Lu CR, Zheng HK, Huang SW, Lin T, Zhu YX, Li FG. 2018. Resequencing of 243 diploid cotton accessions based on an updated A genome identifies the genetic basis of key agronomic traits. *Nature Genetics* **50**:796-802 DOI 10.1038/s41588-018-0116-x.

EI-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M, Richardson L, Salazar GA, Smart A, Sonnhammer EL, Hirsh L, Paladin L., Piovesan D, Tosatto SC, Finn RD. 2019. The Pfam protein families database in 2019. *Nucleic Acids Research* **47**:427-432 DOI 10.1093/nar/gky995.

Eriksson ME, Hanano S, Southern MM, Hall A, Millar, AJ. 2003. Response regulator homologues have complementary, light dependent functions in the Arabidopsis circadian clock. *Planta* **218**:159-162 DOI 10.1007/s00425-003-1106-4.

Farre EM, Kay SA. 2007. PRR7 protein levels are regulated by light and the circadian clock in *Ara bidopsis*.

Plant Journal **52**:548-560.DOI 10.1111/j.1365-313x.2007.03258.x

Flowers TJ. 2004. Improving crop salt tolerance. *Journal of experimental botany* **55**:307-319 DOI 10.1093/jxb/erh003.

Gendron JM, Pruneda-Paz JL, Doherty CJ, Gross AM, Kang SE, Kay SA. 2012. Arabidopsis circadian clock protein, TOC1, is a DNA-binding transcription factor. *Proceedings of the National Academy of Sciences of the United States of America* **109**:3167-3172 DOI 10.1073/pnas.1200355109.

Gome JM, Rodrigues FA, Pagliarini RF, Bendix C, Nakayama TJ, Celaya B, Molinari HBC, Oliveira MCN, Harmon FG, Nepomuceno A. 2014. Diurnal oscillations of soybean circadian clock and drought responsive genes. *PLoS ONE* **9**:e86402 DOI 10.1371/journal.pone.0086402.

Gould PD, Locke JCW, Larue C, Southern MM, Davis SJ, Hanano S, Moyle R, Milich R, Putterill J, Millar AJ, Hall A. 2006. The Molecular Basis of Temperature Compensation in the *Arabidopsis* Circadian Clock. *The Plant Cell* **18**:1177-1187 DOI 10.1105/tpc.105.039990.

Guo AY, Zhu QH, Chen X, Luo JC. 2007. GSDS: A gene structure display server. *Hereditas* **29**:1023-1026 DOI 10.1360/yc-007-1023.

Harmer SL. 2009. The circadian system in higher plants. *Annual Review of Plant Biology* **60**:357-377 DOI 10.1146/annurev.arplant.043008.092054.

Hayama R, Krebs LS, Richter R, Fernández V, Jang S, Coupland G. 2017. Pseudo response reg- ulators stabilize CONSTANS protein to promote flowering in response to day length. *The EMBO Journal* **36**:904-918 DOI 10.15252/embj.201693907.

Hsu PY, Harmer SL.2014. Wheels within wheels: the plant circadian system. *Trends in Plant Science* **19**:240-249 DOI 10.1016/j.tplants.2013.11.007.

Hu Y, Chen JD, Fang L, Zhang ZY, Ma W, Niu YC, Ju LZ, Deng JQ, Zhao T, Lian JM, Baruch K, Fang DD, Liu X, Ruan YL, Rahman M, Han JL, Wang K, Wang Q, Wu HT, Mei GF, Zang YH, Han ZG, Xu CY, Shen WJ, Yang DF, Si ZF, Dai F, Zou LF, Huang F, Bai YL, Zhang YG, Brodt A, Benhamo H, Zhu XF, Zhou BL, Guan XY, Zhu SJ, Chen XY, Zhang TZ. 2019. *Gossypium barbadense* and *Gossypium hirsutum* genomes provide insights into the origin and evolution of allotetraploid cotton. *Nature Genetics* **51**:739-748 DOI 10.1038/s41588-019-0371-5.

Huang G, Wu Z, Percy RG, Bai M, Li Y, Frelichowski JE, Hu J, Wang K, Yu JZ, Zhu Y.2020. Genome sequence of *Gossypium herbaceum* and genome updates of *Gossypium arboreum* and *Gossypium hirsutum* provide insights into cotton A-genome evolution. *Nature Genetics* **52**:516-524 DOI 10.1038/s41588-020-0607-4.

Huang W, Pérez-García P, Pokhilko A, Millar A J, Antoshechkin I, Riechmann JL, Mas P. 2012. Mapping the core of the *Arabidopsis* circadian clock defines the network structure of the oscillator. *Science* **336**:75-79 DOI 10.1126/science.1219075.

Initiative AG. 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**:796-815 DOI 10.1038/35048692.

Ito S, Kawamura H, Niwa Y, Nakamichi N, Yamashino T, Mizuno T. 2009. A genetic study of the Arabidopsis circadian clock with reference to the *TIMING OF CAB EXPRESSION 1* (*TOC1*) gene. *Plant and Cell Physiology* **50**:290-303 DOI 10.1093/pcp/pcn198.

Jackson S, Chen ZJ. 2010. Genomic and expression plasticity of polyploidy. *Current Opinion in Plant Biology* **13**:153-159 DOI 10.1016/j.pbi.2009.11.004.

Jang S , Marchal V, Panigrahi KCS, Wenkel S, Soppe W, Deng XW, Valverde F, Coupland G. 2008. Arabidopsis COP1 shapes the temporal pattern of CO accumulation conferring a photoperiodic flowering response. *The EMBO Journal* **27**:1277-1288 DOI 10.1038/emboj.2008.68.

Kaczorowski KA, Quail PH. 2003. Arabidopsis PSEUDORESPONSEREGULATOR7 is a signaling intermediate in phytochrome regulated seedling deetiolation and phasing of the circadian clock. *Plant Cell* **15**:2654-2665 DOI 10.1105/tpc.015065.

Kahle D. 2013. mpoly: Multivariate Polynomials in R. *The R journal* **5**:181-187 DOI 10.32614/RJ-2013-015.

Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M. 2014. Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Research* **42**:199 DOI 10.1093/nar/gkt1076.

Keily J, Mac Gregor DR, Smith RW, Millar AJ, Halliday KJ, Penfield S. 2013. Model selection reveals control of cold signaling by evening-phased components of the plant circadian clock. *Plant Journal* **76**:247–257 DOI 10.1111/tpj.12303.

Kobayashi Y, Kaya H, Goto K, Iwabuchi M, Araki T.1999. A pair of related genes with antagonis tic roles in mediating flowering signals. *Science* **286**:1960-1962 DOI 10.1126/science.286.5446.1960.

Korneli C, Danisman S, Staiger D. 2014. Differential control of pre-invasive and post-invasive anti- bacterial defense by the Arabidopsis circadian clock. *Plant and Cell Physiology* **55**:1613-1622 DOI 10.1093/pcp/pcu092.

Lee C, Shibata Y, Rao BS, Strahl BD, Lieb JD. 2005. Plant circadian clocks increase photosynthesis, growth, survival, and competitive advantage. *Science* **309**:630-633 DOI 10.1126/science.1115581.

Legnaioli T, Cuevas J, Mas P. 2009.TOC1 functions as a molecular switch connecting the circadian clock with plant responses to drought. The *EMBO Journal* **28**:3745-3757 DOI 10.1038/emboj.2009.297.

Letunic I, Bork P.2019. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Research* **47**:256 DOI 10.1093/nar/gkz239.

LetunicI, Goodstadt L, Dickens NJ, Doerks T, Schultz J,Mott R, Ciccarelli FD, Copley RR, Ponting CP, Bork P. 2002. Recent improvements to the SMART domain-based sequence annotation resource. *Nucleic Acids Research* **30**:242-244 DOI 10.1093/nar/30.1.242.

Li F, Fan G, Lu C, Xiao G, Zou C, Kohel RJ, Ma Z, Shang H, Ma X, Wu J, Liang X, Huang G, Percy RG, Liu K, Yang W, Chen W, Du X, Shi C, Yuan Y, Ye W, Liu X, Zhang X, Liu W, Wei H, Wei S, Huang G, Zhang X, Zhu S, Zhang H, Sun F, Wang X, Liang J, Wang J, He Q, Huang L, Wang J, Cui J, Song G, Wang K, Xu X, Yu JZ, Zhu Y, Yu S. 2015. Genome sequence of cultivated Upland cotton (*Gossypium hirsutum* TM-1) provides insights into genome evolution. *Nature Biotechnology* **33**:524-530 DOI 10.1038/nbt.3208.

Li L, Zhao S, Su J, Fan S, Pang C, Wei H, Wang H, Gu L, Zhang C, Liu G, Yu D, Liu Q, Zhang X, Yu S. 2017. High-density genetic linkage map construction by F2 populations and QTL analysis of early-maturity traits in upland cotton (*Gossypium hirsutum* L.). *PLoS ONE* **12**:e0182918 DOI 10.1371/journal.pone.0182918.

Liu H, Wang HG, Gao PF, Xü JH , Xü TD, Wang JS, Wang BL, Lin CT, Yong FF. 2009. Analy- sis of clock gene homologs using unifoliolates as target organs in soybean (*Glycine max*). *Journal of Plant*

Physiology **166**:278-289 DOI 10.1016/j.jplph.2008.06.003.

Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantita- tive PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**:402-408 DOI 10.1006/meth.2001.1262.

Lu TT, Zhang GF, Sun LR, Wang J, Hao FS. 2017. Genome-wide identification of CBL family and expression analysis of CBLs in response to potassium deficiency in cotton. *PeerJ* ***5***:e3653 DOI 10.7717/peerj.3653.

Más P, Kim WY, Somers DE, Kay KA. 2003. Targeted degradation of TOC1 by ZTL modulates circadian function in *Arabidopsis thaliana*. *Nature* **426**:567-570 DOI 10.1038/nature02163.

Matsushika A, Makino S, Kojima M, Mizuno T. 2000. Circadian waves of expression of the APRR1/TOC1 family of pseudo-response regulators in *Arabidopsis thaliana*: insight into the plant circadian clock. *Plant and Cell Physiology* **41**:1002-1012 DOI 10.1093/pcp/pcd043.

Mcclung CR. 2006. Plant Circadian Rhythms. *The Plant Cell* **18**(4):792-803 DOI 10.1105/tpc.106.040980.

Mizuno T, Nakamichi N. 2005. Pseudo-Response Regulators (PRRs) or True Oscillator Components (TOCs), *Plant and Cell Physiology* **46**:677-685 DOI 10.1093/pcp/pci087.

Murakami M, Matsushika A, Ashikari M, Yamashino T, Mizuno T. 2005. Circadian-associated rice pseudo response regulators (OsPRRs): insight into the control of flowering time. *Bioscience, Biotechnology, and Biochemistry* **69**:410-414 DOI 10.1271/bbb.69.410.

Nakahira Y, Baba K, Yoneda A, Shiina T, Toyoshima Y. 1998. Circadian-regulated transcription of the psbD light-responsive promoter (psbD LRP) in wheat chloroplasts. *Plant Physiology* **118**:1079-1088 DOI 10.1007/978-94-011-3953-3_664.

Nakamichi N, Kusano M, Fukushima A, Kita M, Ito S, Yamashino T, Saito K, Sakakibara H, MizunoT.2009. Transcript profiling of an Arabidopsis *PSEUDO RESPONSE REGULATOR* arrhythmic triple mutant reveals a role for the circadian clock in cold stress response. *Plant and Cell Physiology* **50**:447-462 DOI 10.1093/pcp/pcp004.

Nakamichi N, Kiba T, Henriques R, Mizuno T, Chua NH, Sakakibara H. 2010. PSEU DO-RESPONSE REGULATORS 9, 7, and 5 are transcriptional repressors in the Arabidopsis circadian clock. *Plant Cell* **22**:594-605 DOI 10.1105/tpc.109.072892.

Nakamichi N, Kudo T, Makita N, Kiba T, Kinoshita T, Sakakibara H. 2020. Flowering time con- trol in rice by introducing Arabidopsis clock-associated PSEUDO-RESPONSE REGULATOR 5. *Biosci Biotechnol Biochem* **84**:970-979 DOI 10.1080/09168451.2020.1719822.

Nakamichi N. 2015. Adaptation to the local environment by modifications of the photoperiod re- sponse in crops. *Plant and Cell Physiology* **56**:594-604 DOI 10.1093/pcp/pcu181.

Norihito N, Masanori K, Kanae N, Shogo I, Takafumi Y, Tsuyoshi M, Takeshi M. 2007. Ara- bidopsis clock-associated Pseudo-Response Regulators PRR9, PRR7 and PRR5 coordinately and positively regulate flowering time through the canonical CONSTANS-dependent photoperiodic pathway. *Plant and Cell Physiology* **48**:822-832 DOI 10.1093/pcp/pcm056.

Pokhilko A, Fernández AP, Edwards KD, Southern MM, Halliday KJ, Millar AJ.2012. The clock gene circuit in Arabidopsis includes a repressilator with additional feedback loops. *Molecular Systems Biology* **8**:574-574 DOI 10.1038/msb.2012.6.

Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD. 2018. HMMER web server: 2018 update.

Nucleic Acids Research **46**:200 DOI 10.1093/nar/gky448.

Salome PA, McClung CR. 2005. PSEUDO-RESPONSE REGULATOR 7 and 9 are partially redundant genes essential for the temperature responsiveness of the Arabidopsis circadian clock. *Plant Cell* **17**:791-803 DOI 10.1105/tpc.104.029504.

Seaton DD, Smith RW, Song YH. 2015. Linked circadian outputs control elongation growth and flowering in response to photoperiod and temperature. *Molecular Systems Biology* **11**:776 DOI 10.15252/msb.20145766.

Song YH, Ito S, Imaizumi T. 2013. Flowering time regulation: photoperiod and temperature-sensing in leaves. *Plant Science* **18**:575-583 DOI 10.1016/j.tplants.2013.05.003.

Song YH, Shim JS, Kinmonth-Schultz HA, Imaizumi T. 2015. Photoperiodic flowering: time measurement mechanisms in leaves. *Annual Review of Plant Biology* **66**:441-464 DOI 10.1146/annurev-arplant-043014-115555.

Soy J, Leivar P, González-Schain N, Martín G, Diaz C, Sentandreu M, Al-Sady B, Quail PH, Monte E.2016. Molecular convergence of clock and photosensory pathways through PIF3-TOC1 interaction and co-occupancy of target promoters. *Proceedings of the National Academy of Sciences of the United States of America* **113**:4870-4875 DOI 10.1073/pnas.1603745113.

Strayer C, Oyama T , Schultz TF, Raman R, Somers DE, Más P, Panda S, Kreps JA, Kay SA. 2000. Cloning of the Arabidopsis clock gene TOC1, an autoregulatory response regulator homolog. *Science* **289**:768-771 DOI 10.1126/science.289.5480.768.

Sudhir K, Glen S, Koichiro T. 2016. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* **33**: 1870-1874 DOI 10.1093/molbev/msw054.

Takata N , Saito S, Saito CK , Nanjo T , Shinohara K, Uemura M. 2009. Molecular phylogeny and expression of poplar circadian clock genes, LHY1 and LHY2. *New Phytologist* **18**:808-819 DOI 10.1111/j.1469-8137.2008.02714.x.

Turner A, Beales J, Faure S, Dunford R, Laurie D.2005. The pseudo-response regulator Ppd-H1 provides adaptation to photoperiod in Barley. *Science* **310**:1031-1034 DOI 10.1126/science.1117619.

Uehara TN, Mizutani Y, Kuwata K, Hirota T, Sato A, Mizoi J, Takao S, Matsuo H, Suzuki T, Ito S, Saito AN, Nishiwakiohkawa T, Yamaguchishinozaki K, Yoshimura T, Kay AS, Itami K, Kinoshita T, Yamaguchi J, Nakamichi N.2019. Casein kinase 1 family regulates PRR5 and TOC1 in the Arabidopsis circadian clock. *Proceedings of the National Academy of Sciences of the United States of America* **116**:11528-11536 DOI 10.1073/pnas.1903357116.

Wang KB,Wang ZW, Li FG, Ye WW, Wang JY, Song GL, Yue Z, Cong L, Shang HH, Zhu SL, Zou CS, Li Q, Yuan YL, Lu CR, Wei HL, Gou CY, Zheng ZQ, Yin Y, Zhang XY, Liu K, Wang B, Song C, Shi N, Kohel RJ, Percy RG, Yu JZ, Zhu YX, Wang J, Yu SX. 2012. The draft genome of a diploid cotton *Gossypium raimondii*. *Nature Genetics* **44**:1098-1103 DOI 10.1038/ng.2371.

Wang L, Kim J, Somers DE. 2013. Transcriptional corepressor TOPLESS complexes with pseudo-response regulator proteins and histone deacetylases to regulate circadian transcription. *Proceedings of the National Academy of Sciences of the United States of America* **110**(2): 761-766. DOI 10.1073/pnas.1215010110.

Wang MJ, Tu LL, Yuan DJ, Zhu D, Shen C, Li JY, Liu FY, Pei LL, Wang PC, Zhao GN, Ye ZX, Huang H, Yan FL, Ma YZ, Zhang L, Liu M, You JQ, Yang YC, Liu ZP, Huang F, Li BQ, Qiu P, Zhang QH,

Zhu LF, Jin SX, Yang XY, Min L, Li GL, Chen LL, Zheng HK, Lindsey K, Lin ZX, Udall JA, Zhang XL. 2018. Reference genome sequences of two cultivated allotetraploid cottons, *Gossypium hirsutum* and *Gossypium barbadense*. *Nature Genetics* **51**:224-229 DOI 10.1038/s41588-018-0282-x.

Wang YP, Tang HB, Debarry JD, Tan X, Li JP, Wang XY, Lee TH, Jin HZ, Marler B, Guo H, Kissinger JC, Paterson AH. 2012. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Research* **40**:49 DOI 10.1093/nar/gkr1293.

Wickland DP, Hanzawa Y. 2015. The *FLOWERING LOCUS T/TERMINAL FLOWER 1* gene family: functional evolution and molecular mechanisms. *Molecular Plant* **8**:983-997 DOI 10.1016/j.molp.2015.01.007.

Xie C, Mao XZ, Huang JJ, Ding Y, Wu JM, Dong S, Kong L, Gao G, Li CY, Wei LP. 2011. KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Research* **39**:316-322 DOI 10.1093/nar/gkr483.

Xue ZG, Zhang XM, Lei CF, Chen XJ, Fu YF. 2012. Molecular cloning and functional analysis of one *ZEITLUPE* homolog *GmZTL3* in soybean. *Molecular Biology Reports* **39**:1411-1418 DOI 10.1007/s11033-011-0875-2.

Yamamoto Y, Sato E, Shimizu T, Nakamich N, Sato S, Kato T, Tabata S, Nagatani A, Yam- ashino T, Mizuno T. 2003. Comparative genetic studies on the APRR5 and APRR7 genes belonging to the APRR1/TOC1 quintet implicated in circadian rhythm, control of flowering time, and early photomorphogenesis. *Plant Cell Physiology* **44**:1119-1130 DOI 10.1093/pcp/pcg148.

Yang CP, Tan YR, Yan BY, Gong XX, Wang D, Gao X, Zhang H, Wang P, Li SJ, Wang Y, Zhou LY, Pan YW, Liu JP. 2020. Molecular characterization of clock-associated PSEUDO-RESPONSE REGULATOR 9 gene from Oncidium 'Gower Ramsey'. *Plant Growth Regulation*. **Suppl 1**:1-11DOI 10.1007/s10725-020-00611-6.

Yang Y, Peng Q, Chen GX, Li XH, Wu CY. 2013. OsELF3 is involved in circadian clock regulation for promoting flowering under long-day conditions in rice. *Molecular Plant* **6**:202-217 DOI 10.1093/mp/sss062.

Ye WW, Wang YC, Dou DL. 2010. SeqHunter: a bioinformatics toolbox for local Blast and sequence analysis. *China Journal of Bioinformatics* **8**:364-363 DOI 10.3724/SP.J.1187.2010.00953.

Yu J, Wang J, Lin W, Li SG, Li H, Zhou J, Ni PX, Dong W, Hu SN, Zeng CQ, Zhang JG, Zhang Y, Li RQ, Xu ZY, Li ST, Li XR, Zheng HK, Cong LJ, Lin L, Yin JM, Geng JN, Li GY, Shi JP, Liu J, Lv H, Li J, Wang J, Deng YJ, Ran LH, Shi XL, Wang XY, Wu QF, Li CF, Ren XY, Wang JQ, Wang XL, Li DW, Liu DY, Zhang XW, Ji ZD, Zhao WM, Sun YQ, Zhang ZP, Bao JY, Han YJ, Dong LL, Ji J, Chen P, Wu S, Liu JS, Xiao Y, Bu DB, Tan JL, Yang L, Ye C, Zhang JF, Xu JY, Zhou Y, Yu YP, Zhang B, Zhuang SL, Wei HB, Liu B, Lei M, Yu H, Li YZ, Xu H, Wei SL, He XM, Fang LJ, Zhang ZJ, Zhang YZ, Huang XG, Su ZX, Tong W, Li JH, Tong ZZ, Li SL, Ye J, Wang LS, Fang L, Lei TT, Chen C, Chen H, Xu Z, Li HH, Huang HY, Zhang F, Xu HY, Li N, Zhao CF, Li ST, Dong LJ, Huang YQ, Li L, Xi Y, Qi QH, Li WJ, Zhang B, Hu W, Zhang YL, Tian XJ, Jiao YZ, Liang XH, Jin J, Gao L, Zheng WM, Hao BL, Liu SQ, Wang W, Yuan LP, Cao ML, McDermott J, Samudrala R, Wang J, Wong GK, Yang HM. 2005. The genomes of *Oryza sativa*: A history of duplications. *PLoS Biology* **3**: 38 DOI 10.1371/journal.pbio.0030038.

Yuan DJ, Tang ZH, Wang MJ, Gao WH, Tu LL, Jin X, Chen LL, He YH, Zhang L, Zhu LF, Li Y,

Liang QQ, Lin ZX, Yang XY, Liu N, Jin SX, Lei Y, Ding YH, Li GL, Ruan XA, Ruan YJ, Zhang XL. 2015. The genome sequence of Sea-Island cotton (Gossypium barbadense) provides insights into the allopolyploidization and development of superior spinnable fibre. *Scientific Reports* **5**:17662-17687 DOI 10.1038/srep17662.

Zhang C, Xie QG, Anderson RG, Ng G, Seitz NC, Peterson T, McClung CR, McDowell JM, Kong DD, Kwak JM, Lu H. 2013. Molecular basis of crosstalk between the circadian clock and innate immunity in Arabidopsis. *PLoS Pathogens* **9**:e1003370 DOI 10.1371/journal.ppat.1003370.

Zhang GF, Yue CM, Lu TT, Sun LR, Hao FS. 2020. Genome-wide identification and expression analysis of NADPH oxidase genes in response to ABA and abiotic stresses, and in fibre formation in Gossypium. *PeerJ* **8**:e8404 DOI 10.7717/peerj.8404

Zhang P, Fan SL, Song MZ, Pang CY, Wei HL, Yu SX. 2016. Cloning and Functional Analysis of the Flowering-related Gene GhFLP1 from Upland Cotton (*Gossypium hirsutum* L.). *Cotton Science* **28**: 199-207 DOI 10.11963/issn.1002-7807.201603002.

Zhang TZ , HuY, Jiang WK, Fang L, Guan XY, Chen JD, Zhang JB, Saski CA, Scheffler BE, Stelly DM, Kemp AMH, Wan Q, Liu BL, Liu CX, Wang S, Pan MQ, Wang YK, Wang DW, Ye WX, Chang LJ, Zhang WP, Song QX, Kirkbride RC, Chen XY, Dennis E, Llewellyn DJ, Peterson DG, Thaxton P, Jones DC, Wang Q, Xu XY, Zhang H, Wu HT, Zhou L, Mei GF, Chen SQ, Tian Y, Xiang D, Li XH, Ding J, Zuo QY, Tao L, Liu YC, Li J, Lin Y, Hui YY, Cao ZS, Cai CP, Zhu XF, Jiang Z, Zhou BL, Guo WZ, Li RQ, Chen ZJ. 2015. Sequencing of allotetraploid cotton (*Gossypium hirsutum* L. acc. TM-1) provides a resource for fiber improvement. *Nature Biotechnology* **33**:531-537 DOI 10.1038/nbt.3207.

Zhang Y, Wang XF, Ding ZG, Ma Q, Zhang GR, Zhang SL, Li ZK, Wu Q, Zhang GY,Ma ZY. 2013. Transcriptome profiling of *Gossypium barbadense* inoculated with *Verticillium dahliae* provides a resource for cotton improvement. *BMC Genomics* **14**:637 DOI 10.1186/1471-2164-14-637.

Zhu JY, Oh E, Wang TN, Wang ZY. 2016. TOC1-PIF4 interaction mediates the circadian gating of thermores-ponsive growth in Arabidopsis. *Nature Communications* **7**:13692 DOI 10.1038/ncomms13692.

Figure legends

Figure 1 Phylogenetic and collinearity analysis of PRR proteins in cotton.

(A) Phylogenetic tree of the *PRR* gene family; (B) Gene duplication and collinearity analysis among cotton *PRR* genes (green lines and brown indicates paralogous genes in *G. hirsutum* and *G. barbadense*, orange lines indicates orthologous genes between *G. arboreum* and *G. hirsutum*, black lines indicates orthologous genes between *G. arboreum* and *G. raimondii*, blue lines indicates orthologous genes between *G. barbadense* and *G. hirsutum*, seagreen indicates orthologous genes between *G. arboreum* and *G. barbadense*, lightsteelblue indicates orthologous genes between *G. barbadense* and *G. raimondii*, yellow indicates orthologous genes between *G. raimondii* and *G. hirsutum*).

Figure 2 Genetic structure and motif prediction of PRR members.

(A) Genetic structure of *GhPRR* genes and motif prediction of GhPRR proteins; (B) Cis-elements prediction of *GhPRR* promoters.

Figure 3 Periodic xpression pattern of PRR gene family and related genes during 24 hours in LMY19.

Figure 4 Expression analysis and KEGG enrichment of *PRR* genes.

(A)Differential expression genes analysis; (B) KEGG pathway enrichment of DEGs PEG_6 h group;

(C)Expression pattern of *GhPRR* genes with PEG treated at 1, 3 and 6 h.

DEGs: differentially expressed genes.

Figure 5 qRT-PCR analysis of *PRR* genes under PEG treatment at 1, 3, 6 h. The different lowercase letters indicate significant differences at $p<0.05$. Differences analysis were compared using one-way ANOVA and Duncan test.

Supplemental Information

Figure S1 Distributions of the PRR family genes on chromosomes in *Gossypium spp*. (A) Distributions of *GaPRR* genes on chromosomes in *G. arboretum*; (B) Distributions of *GbPRR* genes on chromosomes in *G. barbadense*; (C) Distributions of *GhPRR* genes on chromosomes in *G. hirsutum*; (D) Distributions of *GrPRR* genes on chromosomes in *G. raimondii*.

Table S1 qRT-PCR primers for *GhPRR* and related genes used in this study.

Table S2 Rename information of *PRR* genes.

Table S3 The general information of the PRR gene family.

632 **Table S4 Orthologous relationships among four *Gossypium* species.**

633 **Table S5 The duplication of PRR gene pairs in *Gossypium spp*.**

634 **Table S6 FPKM values of *GhPRR* genes.**

635

636

Figure 1

Phylogenetic and collinearity analysis of PRR proteins in cotton.

(A) Phylogenetic tree of the *PRR* gene family; (B) Gene duplication and collinearity analysis among cotton *PRR* genes (green lines and brown indicates paralogous genes in *G. hirsutum* and *G. barbadense*, orange lines indicates orthologous genes between *G. arboreum* and *G. hirsutum*, black lines indicates orthologous genes between *G. arboreum* and *G. raimondii*, blue lines indicates orthologous genes between *G. barbadense* and *G. hirsutum*, seagreen indicates orthologous genes between *G. arboreum* and *G. barbadense*, lightsteelblue indicates orthologous genes between *G. barbadense* and *G. raimondii*, yellow indicates orthologous genes between *G. raimondii* and *G. hirsutum*).

Figure 1

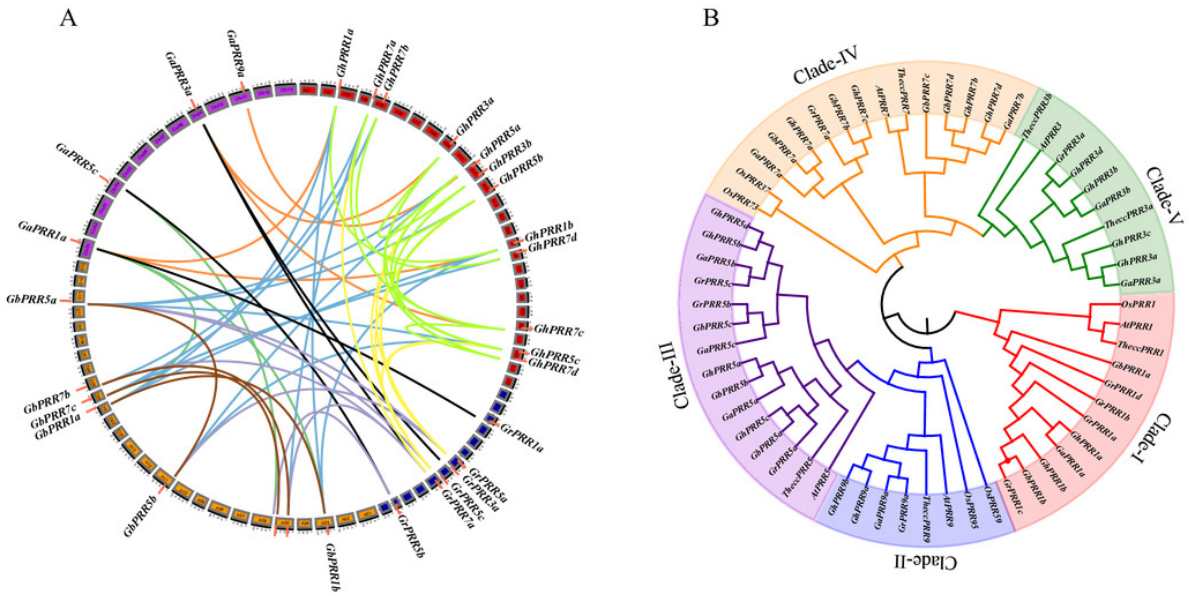


Figure 2

Genetic structure and motif prediction of PRR members.

(A) Genetic structure of *GhPRR* genes and motif prediction of GhPRR proteins; (B) Cis-elements prediction of *GhPRR* promoters.

Figure 2

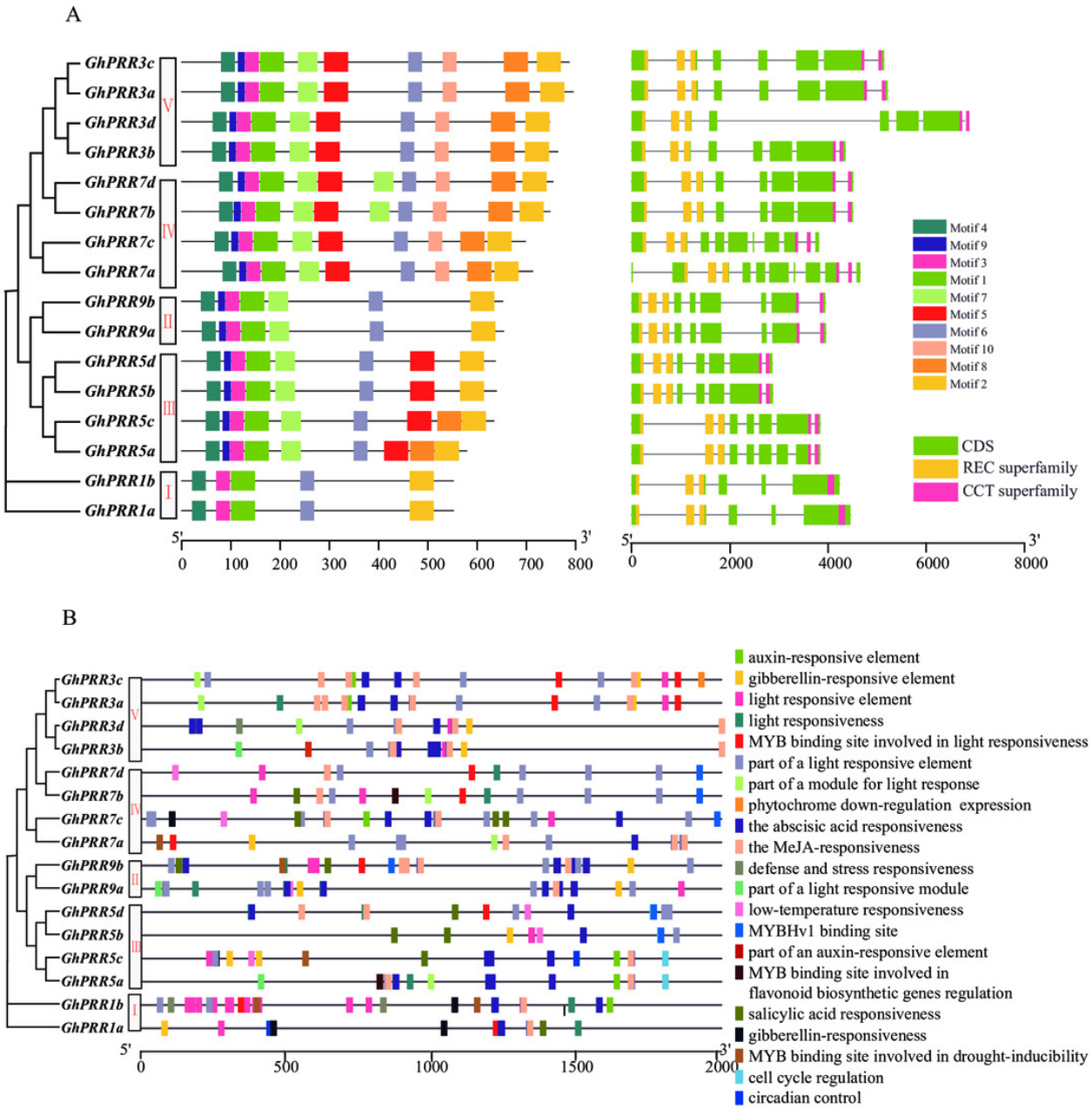

A
GhPRR3c
GhPRR3a
GhPRR3d
GhPRR3b
GhPRR7d
GhPRR7b
GhPRR7c
GhPRR7a
GhPRR9b
GhPRR9a
GhPRR5d
GhPRR5b
GhPRR5c
GhPRR5a
GhPRR1b
GhPRR1a
V
IV
II
III
I
5'
3'
0 100 200 300 400 500 600 700 800
0 2000 4000 6000 8000
Motif 4
Motif 9
Motif 3
Motif 1
Motif 7
Motif 5
Motif 6
Motif 10
Motif 8
Motif 2
CDS
REC superfamily
CCT superfamily
B
0 500 1000 1500 2000
auxin-responsive element
gibberellin-responsive element
light responsive element
light responsiveness
MYB binding site involved in light responsiveness
part of a light responsive element
part of a module for light response
phytochrome down-regulation expression
the abscisic acid responsiveness
the MeJA-responsiveness
defense and stress responsiveness
part of a light responsive module
low-temperature responsiveness
MYBHv1 binding site
part of an auxin-responsive element
MYB binding site involved in flavonoid biosynthetic genes regulation
salicylic acid responsiveness
gibberellin-responsiveness
MYB binding site involved in drought-inducibility
cell cycle regulation
circadian control

Figure 3

Periodic xpression pattern of PRR gene family and related genes during 24 hours in LMY19.

Figure 3

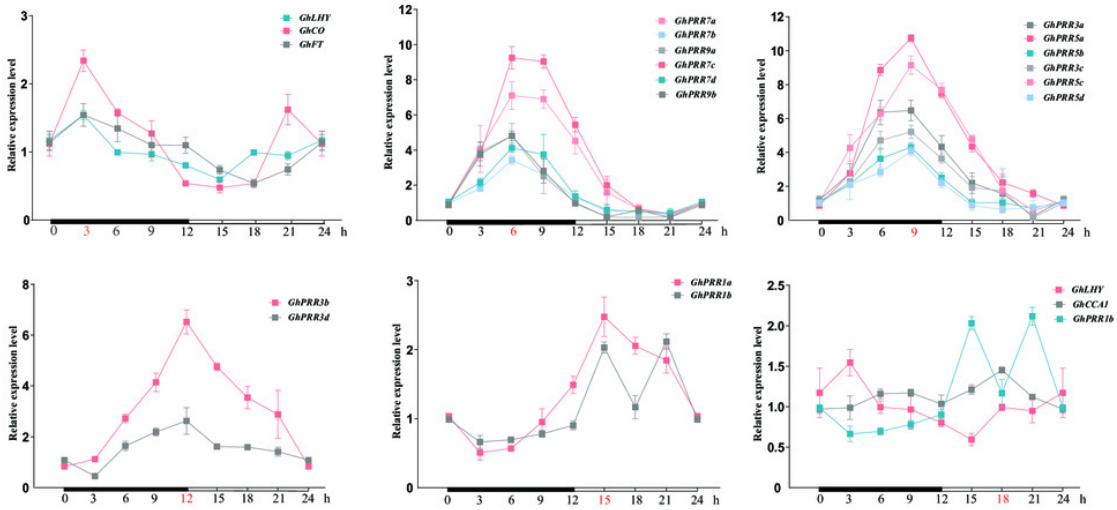

Relative expression level
GhLHY
GhCO
GhFT
GhPRR7a
GhPRR7b
GhPRR9a
GhPRR7c
GhPRR7d
GhPRR9b
GhPRR3a
GhPRR5a
GhPRR5b
GhPRR3c
GhPRR5c
GhPRR5d
GhPRR3b
GhPRR3d
GhPRR1a
GhPRR1b
GhLHY
GhCCA1
GhPRR1b
0 3 6 9 12 15 18 21 24 h

Figure 4

Expression analysis and KEGG enrichment of *PRR* genes.

(A)Differential expression genes analysis; (B) KEGG pathway enrichment of DEGs PEG_6 h group; (C)Expression pattern of *GhPRR* genes with PEG treated at 1, 3 and 6 h.

Figure 4

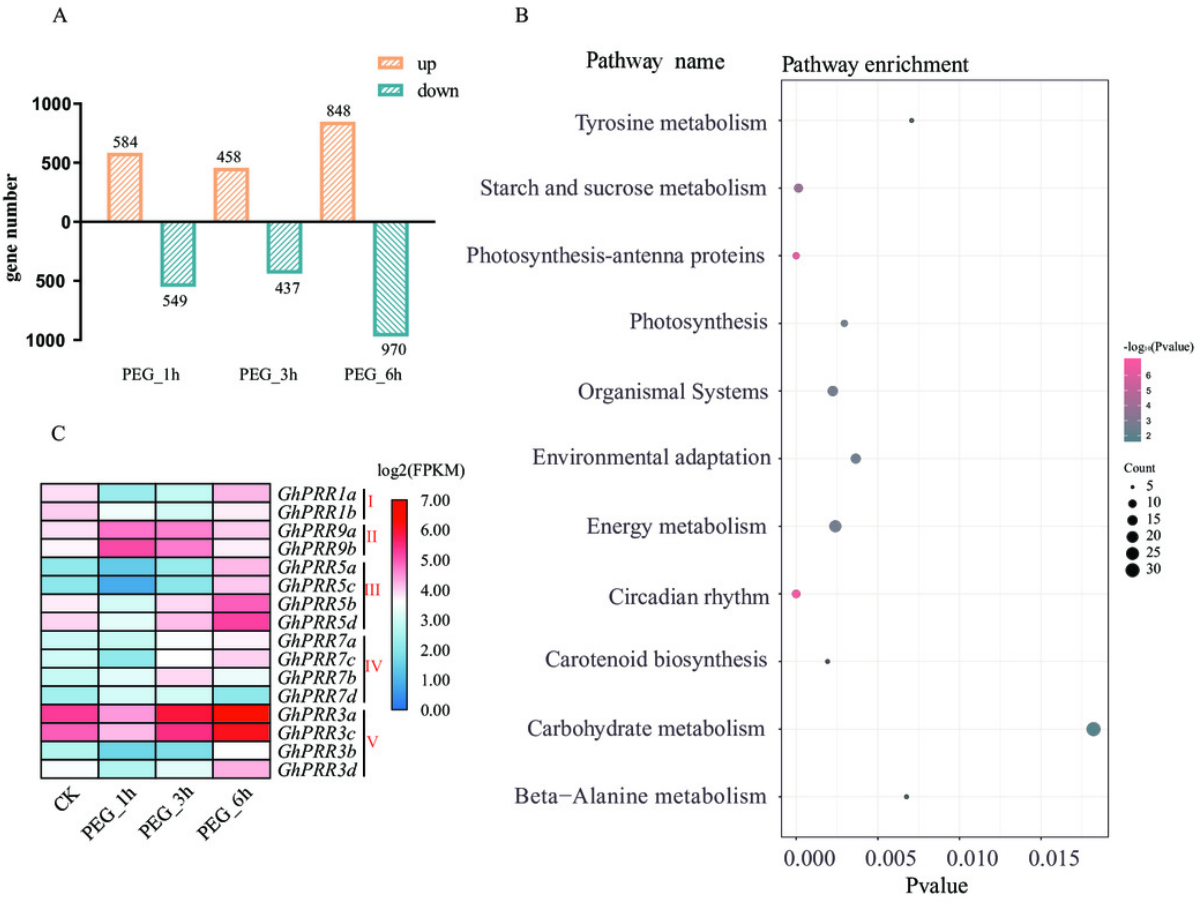
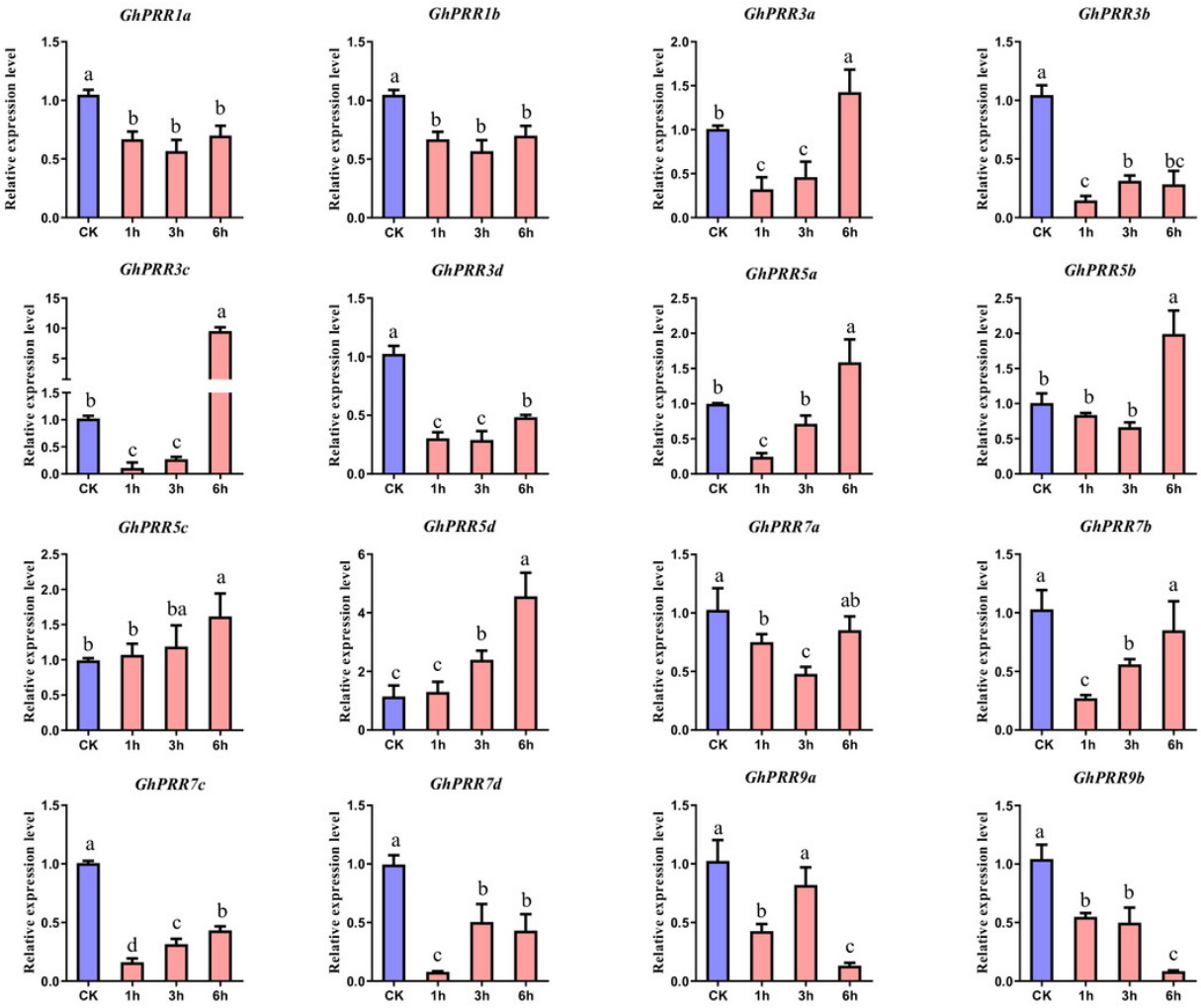

A
up
down
gene number
1000
500
0
500
1000
584
549
458
437
848
970
PEG_1h
PEG_3h
PEG_6h
B
Pathway name
Pathway enrichment
Tyrosine metabolism
Starch and sucrose metabolism
Photosynthesis-antenna proteins
Photosynthesis
Organismal Systems
Environmental adaptation
Energy metabolism
Circadian rhythm
Carotenoid biosynthesis
Carbohydrate metabolism
Beta−Alanine metabolism
0.000
0.005
0.010
0.015
Pvalue
-log10(Pvalue)
Count
5
10
15
20
25
30
C
log2(FPKM)
7.00
6.00
5.00
4.00
3.00
2.00
1.00
0.00
GhPRR1a
GhPRR1b
GhPRR9a
GhPRR9b
GhPRR5a
GhPRR5c
GhPRR5b
GhPRR5d
GhPRR7a
GhPRR7c
GhPRR7b
GhPRR7d
GhPRR3a
GhPRR3c
GhPRR3b
GhPRR3d
I
II
III
IV
V
CK
PEG_1h
PEG_3h
PEG_6h

Figure 5

qRT-PCR analysis of *PRR* genes under PEG treatment at 1, 3, 6 h.

The different lowercase letters indicate significant differences at $p<0.05$. Differences analysis were compared using one-way ANOVA and Duncan test.

Figure 5


GhPRR1a
GhPRR1b
GhPRR3a
GhPRR3b
GhPRR3c
GhPRR3d
GhPRR5a
GhPRR5b
GhPRR5c
GhPRR5d
GhPRR7a
GhPRR7b
GhPRR7c
GhPRR7d
GhPRR9a
GhPRR9b
Relative expression level
CK
1h
3h
6h