Genome-wide identification and analysis of WRKY gene family in jute(Corchorus capsularis)

Lilan Zhang
Fujian Agriculture and Forestry University
Xuebei Wan
Fujian Agriculture and Forestry University
Yi Xu
Fujian Agriculture and Forestry University
Sylvain Niyitanga
Fujian Agriculture and Forestry University
Jianmin Qi
Fujian Agriculture and Forestry University
Liwu Zhang (zhang_liwu@hotmail.com)
Fujian Agriculture and Forestry University https://orcid.org/0000-0001-8465-2476

Research article

Keywords: Jute, WRKY transcription factor, GA3 stress, Expression analysis

Posted Date: December 16th, 2019

DOI: https://doi.org/10.21203/rs.2.15299/v2

Version of Record: A version of this preprint was published on September 1st, 2020. See the published version at https://doi.org/10.1186/s12870-020-02617-8.

Abstract

Background WRKY transcription factor is a kind of transcription factor which plays an important role in plant response to biotic, abiotic stress, plant growth and development. However, little information was available about the WRKY genes in jute (Corchorus capsularis).Results In the present study, 43 jute WRKY (CcWRKY) genes were identified by using Pfam database domain search and BLAST homology alignment based on the transcriptome data of jute. And the gene structure, phylogeny, conserved domain and three-dimensional structure of protein were also analyzed by GSDS2.0, MEGA7.0, DNAMAN5.0, WebLogo 3 and SWISS-MODEL bioinformatics tools. According to the WRKY conserved domain features and the evolution analysis with Arabidopsis thaliana , 43 members were divided into three classes: I, II and III containing 9, 28, 6 members, respectively. According to the evolutionary relationship, class II further divided into five subclasses: II-a (2), II-b (7), II-c (7), II-d (6) and II-e (6). Genetic structure analysis showed that exon and intron number of CcWRKY genes had high variability (3-11 exons), even within the same subgroup. Most of the CcWRKY genes were expressed in different tissues, but they were mainly expressed in stem bark and stem stick. After GA 3 stress, the expression of most WRKY genes in GA 3 -sensitive variety "Aidianyehuangma" was significantly different from that of normal variety "Huangma 179". These results indicated that CcWRKY genes play an important role in gibberellin biosynthesis pathway and fiber development.Conclusions CcWRKY proteins are highly conserved, the length of the gene sequence and the number of introns varied widely, all WRKY genes showed a variety of expression patterns in different tissues, most of the WRKY genes responded to GA 3 stress, which play an important role in gibberellin biosynthesis pathway and bast fiber development.

Background

The WRKY gene family is a transcription factor that exists only in plants. It is mainly involved in transcriptional regulation and signal transduction processes in plants[1].In the transcriptional regulatory network, WRKY transcription factors bind specific DNA sequences to activate or repress transcription of multiple target genes[2, 3].The conserved WRKY domain contains approximately 60 amino acid residues. In the WRKY domain, a conserved WRKYGQK hexapeptide sequence is usually followed by a C2H2- or C2HC-type zinc finger structure. According to the number of WRKY domains and the type of zinc finger structure, the WRKY family can be divided into 3 groups: Group Ⅰ, Group Ⅱ and Group Ⅲ[4]. Group II could be further divided into five subgroups: II-a, II-b, II-c, II-d and II-e. Group Ⅰ contains two WRKY domains and C_2H_2 zinc finger structure, Group Ⅱ contains a WRKY domain and a C_2H_2 zinc finger structure, and Group Ⅲ contains a WRKY domain and a C_2-H-C zinc finger structure[5].

Since the first WRKY gene (*SPF1*) was cloned from sweet potato[6], WRKY genes have been found in many plants, such as *Arabidopsis* [7], rice [8], barley [9], rapeseed [10], soybean [11], corn [12], potato [13], tobacco [14], poplar [15], cotton [16], cucumber [5] and grape [17]. At present, WRKY transcription factors have been found to play an important role in plant defense response to environmental stresses and growth development. The transcription factor WRKY70 is a common component in salicylic acid (SA)- and jasmonic acid (JA)-mediated signal pathways[18]. Expression of WRKY70 is activated by SA and

repressed by JA [18]. In *Arabidopsis*, some WRKY transcription factors are positive regulators of ABA(Abscisic acid)-mediated stomatal closure and hence drought responses [19]. AtWRKY75 in *Arabidopsis thaliana* and OsWRKY31 in rice are both related to the growth of lateral root [20, 21].The synergistic interaction of ABA-inducible WRKY genes in regulating GAMYB (GA-inducible transcriptional activator)-mediated GA (Gibberellin) signaling in aleurone cells, thereby establishing a novel mechanism for ABA and GA signaling cross-talk[22].

Jute (*Corchoruscapsularis* L) is a diploid (2n=14) natural fiber plant, which belongs to the Malvaceae family [23]. It is mainly cultivated in Bangladesh, India, and China. The bast fiber of jute is mainly used to make fabrics, ropes and threads, which is an indispensable part in human's life. At present, most of the fibrin use, are mainly coming from limited petrochemical products, which are not only expensive, but also harmful to the environment due to the emission of greenhouse gases to the atmosphere. In recent years, a number of people were interested in plant fiber textiles and production technology[24].This was due to the reason that plant fibers are renewable, eco-friendly, and 100% biodegradable. Jute bast fiber has excellent moisture absorption, air permeability and strong antibacterial property, so it has great application value and development potential. In these perspectives, the processing technology and new spinning technology of jute bast fiber were constantly improving and developing, resulting in more and more refined jute products that will hugely promote space in clothing, home textile, building materials and other fields. So far, jute has become one of the most economic fiber crops in the world, known as "gold fiber". All the industries related to jute bast fiber demand sufficient and high-quality fiber. However, both quality and yield of jute are constrained by various stresses attributed to global climate change notably abiotic and biotic stresses.

At present, many studies have shown that WRKY genes plays an important role in plant growth, development, stress tolerance, especially in fiber development[25]. However, information about the WRKY genes in jute is poorly understood. Therefore, this was the first time to identify and analyzed the WRKY transcription factors in jute. The aims of this study were: (1) Identification and structure analysis of WRKY gene family in jute. (2) Phylogenetic analysis and expression analysis of the WRKY gene family in jute after the stressed with exogenous GA_3. (3) the effect of WRKY genes on bast fiber development in jute. Understanding the structure and roles of WRKY gene family can provide basic information for manipulating this gene family in order to increase jute resistance to stresses and fiber yield as well.

Methods

Plant materials

Varieties, "Huangma 179" and "Aidianyehuangma", were used as plant materials in this study. These two varieties of white jute were grown in the farm of Fujian agriculture and forestry university during the growing season in May, 1st, 2017. 10 days after sowing, hypocotyl samples were independently collected from three "Huangma 179" plants. On July 1st, 2017 (60 days after sowing), the stem bark and stem stick were independently collected from three "Huangma 179", stem bark were independently collected from

three "Aidianyehuangma", separately. On September 1st, 2017(120 days after sowing), the stem barks were independently collected from three "Huangma 179", separately.

To identify the expression of CcWRKY genes after GA_3 stress, six jute plants were treated with GA_3 stress 60 days after sowing. The GA_3 stress treatment was carried out for 4 hours and 72 hours, respectively. After the GA_3 stress treatment, the stem barks of each jute were obtained from "Huangma 179" and "Aidianyehuangma". In addition, three jute plants of the same age were sampled separately as control. All samples were immediately frozen in liquid nitrogen and stored in a refrigerator at -80℃ for follow-up analysis. Three samples come from each tissue were stored separately as three biological replicates (Additional file 1: Table S1).

Identification and Analysis of Conservative Domain of CcWRKY

The assembled sequences data have been deposited at the NCBI Sequence Read Archive (SRA, http://www.ncbi.nlm.nih.gov/Traces/sra) vide SRA SRP215917. The assembly of CcWRKY genes and ORF analysis were performed according to Islam et al [23] and Zhang, L. et al.[26].To obtain the WRKY family genes in the jute genome, a local BLASTP search was performed to identify complete WRKY members, using *Arabidopsis* WRKY protein sequences as query sequences. Further, conservative domain prediction software Pfam was used to ensure that all candidate genes contain WRKY conserved domains [27]. The WRKY domain proteins of WRKY genes were identified and analyzed using DNAMAN5.0 software and their conservative structure prediction was performed in Weblogo (http://weblogo.berkeley.edu/logo.cgi). Finally, a total of 43 members of WRKY transcription factor in jute were modeled as well as the tertiary structure of WRKY protein identified using SWISS-MODEL [28],

Phylogenetic analysis of CcWRKY protein

Phylogenetic and molecular evolutionary analysis was conducted using MEGA7 (http://www.megasoftware.net)[29] with pairwise distance and the neighbor-joining algorithm. The p-distance method was used to compute the evolutionary distances, which were used to estimate the number of amino acid substitutions per site. Conducting 1,000 bootstrap sampling steps[12] established the reliability of each tree. Arabidopsis WRKY proteins (AtWRKY) were added to phylogenetic analysis to facilitate the subgroup classification of CcWRKY. WRKY family sequences from *Arabidopsis* were downloaded from TAIR (http://www.arabidopsis.org/).

Expression analysis of CcWRKY genes at different growth stages

The expression analysis CcWRKY genes was performed according to the FRKM value of different genes in different tissues (hypocotyls-10d, leaf-60d, root-60d, stem bark-60d, stem stick-60d, stem bark-120d), the heatmaps was drawn by R language. The data obtained were calculated by using log function.

Effects of GA_3 on the development of bast fiber

To explore the effect of GA_3 on the development of bast fiber in jute, "Huangma 179" was used as the control variety and "Aidianyehuangma" as the test variety. According to the method of MS medium[30], the standard medium (CK) and 0.1 mg·L^{-1} GA_3 (GA) exogenous hormone medium were prepared. The common variety "Huangma 179" and sensitive to gibberellin variety "Aidianyehuangma" were used as experimental materials. Each pair has 3 bottles as biological repeats, 30 seeds were planted in each bottle, and the length of hypocotyl was measured when the first true leaf was grown. Base on homologous protein sequence in Arabidopsis as query sequence, the other fiber marker genes in jute were identified by local BLASTP. According to the FRKM values of fiber marker genes and CcWRKY genes in stem bark (60-day-after-sowing) before and after GA_3 treatment, the histogram was drawn with Excel. The data obtained were calculated by using log function, formula $2^{-\triangle\triangle CT}$ and SPSS. The upstream promoter of CcWRKYs were predicted and analyzed by PlantCARE.

RNA extraction and quantitative real-time PCR (qRT-PCR)

The total RNA was extracted from four different tissues (leaves, roots, stem bark and stem stick) of jute plant at the vigorous growth stage i.e. 60-day-after-sowing using. EZNA Plant RNA Kit (from OMEGA) following manufacturer's instructions. The cDNA was synthesized by a reverse transcription kit (from Takara) according to the manufacture's procedures. This experiment used the Actin[31] as the reference gene by qRT-PCR. qRT-PCR reaction system: 10 μl of GoTaq® qPCR Master Mix, 0.4 μl of left primer(10μM), 0.4 μl of right primer(10μM), 2μl of cDNA, and 7.2μl of Nuclease-Free Water. Amplifications were performed with an initial 10min step of 95℃ followed by 40 denaturation cycles at 95℃ for 15s and primer annealing at 60℃ for 1min. The melting curve used the default program of the instrument (ABI7500). All experiments were used in triplicate for each sample and relative gene expression levels were calculated using the formula $2^{-\triangle\triangle CT}$ as earlier described [32]. All data was analyzed using Microsoft Excel. The primers were designed to avoid the WRKY conserved domain.

Results

Identifcation of full length CcWRKY genes

In this study a local BLASTP search was performed to identify complete WRKY members in jute, using *Arabidopsis* WRKY protein sequences as query sequences and also Pfam was used to detect their conserved domain. As result a total of 43 candidate genes containing WRKY domain (named as CcWRKY) were identified, as shown in Table 1.

Analysis of conservative domain of CcWRKY genes

The conservative domain of CcWRKY gene sequences were identified and analyzed using DNAMAN5.0 software and. conservative structure prediction was performed by Weblogo. The results showed that the conserved domains of WRKY gene family in jute could be divided into three groups: I, II and III. Group I had nine members. It could be further divided into I-C and I-N subgroups. Group I contain two WRKY

domains and zinc finger structures, and the zinc finger structure are $CX_4C_{22-23}HXH$.Group II could be further divided into subgroups II-a, II-b, II-c, II-d and II-e, with 2, 7, 7, 6 and 6 members, respectively. In II-a, II-b, II-d and II-e, the heptapeptide domain and zinc finger structure of WRKY at C-terminal were WRKYGQK and $CX_5C_{23}HXH$, while in II-c, the heptapeptide domain and zinc finger structure of WRKY at C-terminal were WRKYGQK and $CX_4C_{23}HXH$. There were six members in group III. The heptapeptide domain and zinc finger structure of WRKY at C-terminal were WRKYGQK and $CX_7C_{23}HXC$ (Fig1). Moreover, the present results showed that there are still mutations in its protein sequence, though WRKY transcription factor has a much conserved WRKY domain. Among the 43 members of WRKY transcription factor identified in jute, the conserved domain of one gene (WRKYGQK) and the zinc finger structure of four genes were all mutated (Additional file 2: Table S2). This variation indicated that despite the structurally high conserved WRKY gene family, some variations still occur in its WRKY domain, which also illustrated that the plant WRKY gene family had diversity in the evolutionary process.

Phylogenetic analysis of CcWRKY protein in diverse species

BY comparing the known WRKY region of *Arabidopsis thaliana* WRKY protein with CcWRKY, the WRKY domain sequence of CcWRKY protein was clustered and analyzed using MEGA7 (Fig2). These CcWRKY proteins can be divided into three groups: I, II and III. And Group II can be divided into II-a, II-b, II-c, II-d and II-e subgroups. The classifications of phylogenetic tree analysis were consistent with the results of Figure 1(Table 1).

Structure analysis of intron and exon of WRKY in jute

In this study, the number of exons and introns of jute WRKY gene were analyzed and the results are shown in Figure 3. The number of exons varied from 3 to 11.21 WRKYs(48.84%) contained 3 exons, 5 WRKYs(11.63%) contained 4 exons, 8 WRKYs (18.60%) contained 5 exons,6 WRKYs (13.95%) contained 6 exons. From the groups, Group II c+d+e and group III were relatively conservative, while Group I and Group II a+b+c's structures were significantly different and changed greatly. Most CcWRKYs in Group II c+d+e and group III contain 3 exons except Ccv40151700 (4 exons) and Ccv40018590 (4 exons).

Analysis of tertiary structure of protein

The tertiary structure of protein is further coiled and folded on basis of the secondary structure. The tertiary structure of CcWRKY protein was conducted by SWISS-MODEL. The majority of the 43 amino acid sequences have the similar three-dimensional structure. One representative homology modeling from CcWRKY gene family was shown in Figure S1, and consists of several beta folding. Their tertiary structure were quite similar with that of *Arabidopsis thaliana*[33]. It had also proved that the CcWRKY gene family is highly conserved in structure.

Expression analysis of CcWRKY genes in different tissues

Tissue specific expression of genes is often considered as markers of specific gene functions in this tissue. Since WRKY genes are related to the bast fiber development of plants[34, 35], we mainly focus on the expression of CcWRKY genes at different stages of stem growth. Based on the RNA-seq data, we used R language to draw the heatmap of the expression patterns of CcWRKY genes in different stem growth stages (Additional file 4: Fig. S2). The difference of gene expression is generally represented by colors, red represents high expression and blue represents low expression. The results showed that all the CcWRKY genes were expressed in the stem of jute, and the expression of WRKY genes differ at different stem growth stages. Meanwhile, it proved that there were no pseudogenes in 43 genes. From Figure S2, we could see that 43 genes were divided into two categories. The expressions of 13 genes were lower in the different tissues of jute, and the others were higher. Totally, 10 WRKYs were highly expressed in leaf(60d), 3 WRKYs were highly expressed in hypocotyls (10d), 2 WRKYs were highly expressed in stem stick(60d), 2 WRKYs were highly expressed in stem bark(60d), 14 WRKYs were highly expressed in root(60d), and 12 WRKYs were highly expressed in stem bark(120d). It could be seen that the WRKY genes were mainly expressed in the stem bark of jute. With the continuous growth of jute, the bast fiber of jute will gradually accumulate in the stem bark. Therefore, it is believed reasonably that the WRKY genes are involved in bast fiber development in jute. For example, Ccv40032460 was highly expressed in hypocotyls (10d), lowly expressed in stem bark (60d), and no expression in stem bark (120d). It suggests that this gene may play a negative regulatory role in jute fiber accumulation.

GA_3stress analysis of CcWRKY genes involved in cell wall formation

According to our previous research[36], "Aidianyehuangma" is a dwarf variety that sensitive to GA3. "Huangma 179" and "Aidianyehuangma" were planted in 0.1 $mg{\cdot}L^{-1}$ GA_3 exogenous hormone medium (GA) and control group MS medium (CK), respectively, and the length of hypocotyl was measured when the first true leaf was grown. It could be seen that the hypocotyl length of " Aidianyehuangma " treated by GA_3 (4.58cm) is higher than that of CK (1.77cm) (Additional file 5: Fig. S3). The average lengths of hypocotyl of "Aidianyehuangma" and "Huangma 179" treated with GA_3 was 4.58cm and 4.67cm respectively, while that of "Huangma 179" without GA_3 was 4.61cm (Additional file 5: Fig. S3). The results showed that there were no significant differences in the lengths of hypocotyl among the three groups. After stressing the plant with GA_3, it was found that the plant height of "Aidianyehuangma" could be significantly increased, which would greatly improve the fiber yield of the variety. This dwarf variety is very suitable for studying the relationship between GA_3 and fiber development.

To further explore the relationship between fiber development, CcWRKYs and GA_3, we selected some important fiber related genes as marker genes. These genes were include,CesA1 (CesA, Cellulose synthase), CesA4, CesA7, CesA8, CCoAOMT (Caffeioyl coenzyme A methyltransferase), 4CL (4-Coumarate: Coenzyme A Ligase), Ent-copalyl diphosphate synthase, Ent-kaurene oxidase, Ent-kaurene synthase, Ent-kaurenoic acid oxidase, GA 20-oxidase, Gibberellin 2,3-hydroxylase and Gibberellin C13 oxidase. In the vigorous growth period (60 days after sowing), the stem barks were treated with GA_3 stress for "Huangma 179" and " Aidianyehuangma ". Then, the samples were taken after 4 hours and 72

hours, respectively. The samples without GA_3 treatment could be used as control. We analyzed the RNA-seq results of these materials, and then drew the corresponding histogram (Additional file 6: Fig. S4, Additional file 7: Fig. S5), the up column indicated that the gene expressions were up-regulated, and the down column showed the gene expressions were down-regulated. The expression of WRKY genes mostly changed significantly under GA_3 stress (like other fiber related marker genes), especially for the down regulated genes. By comparing the expression of CcWRKY genes under different treatment time (4h and 72h) after spraying GA_3, the expression of most of CcWRKY genes (31 genes) changed in the same trend in "Huangma 179", similar results were found in "Aidianyehuangma".

From these, we also found21 CcWRKY genes and most of the fiber related marker genes of "Aidianyehuangma" were sensitive to the GA_3 stress. The variations of expression of these genes in "Aidianyehuangma" were more significant than those in "Huangma 179" (Fig 4 and Fig 5).This indicated that these CcWRKY genes, were similar to that of like other fiber related marker genes, played a certain role in the increase of fiber yield of "Aidianyehuangma" under GA_3 stress. In addition, the promoters of these CcWRKYs were analyzed by PlantCARE. The results showed that the promoters of most CcWRKY genes contained elements related to gibberellin, such as GARE-motif, P-box and TATC-box (Table 2). The CcWRKY genes responded to the stress of GA_3 and could increase the fiber yield of the "Aidianyehuangma". This suggested that WRKY genes might also be involved in the growth and development of bast fiber like other fiber genes in jute.

To verify the accuracy of the gene expression, 9CcWRKY genes were randomly selected for qRT-PCR analysis (Additional file 8: Fig. S6). The results of qRT-PCR corresponded to the results of FPKM.

Discussion

CcWRKY transcription factors in jute

WRKY transcription factors are one of the largest families of transcriptional regulators in plants. They play important roles in plant growth and development, as well as defensive in biotic and abiotic stresses[1, 37-39].In this study, there were at least 43 WRKY members in jute genome, and the numbers were similar to those of sugar beet(40)[40], canola(43)[41] and castor(47)[42]. However, these numbers were different from those of other plants, such as *Arabidopsis thaliana*(74), cotton(116)[16], soybean(188)[43], poplar(104)[15], tomato(81)[44], cabbage(145)[45], whose members were more than those of jute. By comparing species, the numbers of WRKY genes in different species is not proportional to their genome size. Nowadays, researchers have suggested that gene duplication, segmental duplication and whole genome duplication play important roles in the mass production of gene families[43].Unfortunately, due to the lack of data on jute research, genome data also only published a draft, so many problems failed to get a satisfactory answer. We suspected that jute genome WRKY genes were less than other species, perhaps because they did not experience whole genome replication as other species do. However, with the improvement of genome sequencing and assembly, upgrade and update of

the search and analysis software, discovery of variable splicing in genomes and the continuous advance of jute related research, we believed that the new WRKY members also existed in the genome of jute.

In general, the locations of introns and exons in the genome may provide important evidences for their evolutionary relationships. In this study, we systematically and comprehensively analyzed the distributions and lengths of exons and introns of the members of the WRKY gene family. By analyzing the gene structures of jute WRKY, it was found that its members consisted of 3 to 11 exons, and nearly half of them were 3 exons. These results provided valuable information for the study of jute and the evolutions of the WRKY gene family in other species. Moreover, the members of the WRKY family have similar three-dimensional structures, which were formed by several beta folds. It is similar to the 3D structure of the domains of *Arabidopsis* WRKY protein in the database[33].

CcWRKY genes involved in cell wall formation

Both Jute and cotton are important fibre crops in the world. So, we usually would compare the results of jute with that of cotton. According the expressions analysis of CcWRKY genes in different tissues, we could find that the mostly group III CcWRKY genes (Ccv40018580, Ccv40120290, Ccv40170160 and Ccv40154680) were highly expressed during the fiber development except two genes(Ccv40018590 and Ccv40064890).Theses results in jute were consistent to that in cotton[46].In recent years, the study[47] has shown that *Arabidopsis* WRKY transcription factor was involved in the formation of secondary cell wall of cells, which could significantly increase plant biomass. In this study, most of the CcWRKY genes expressed differently in the stem bark and hypocotyls. This interesting phenomenon leaded us to believe that there was a certain connection between WRKY transcription factors and bast fiber development.

It has long been known that cellulose synthesis in plants was regulated by various phytohormones. GA_3 is an important hormone for plant growth and development throughout the whole life cycle. Therefore in this study, the gene expression patterns of 43 CcWRKY transcription factor genes under GA_3 stress were systematically analyzed. Because of the great influences of GA_3 on plant heights, the dwarf variety "Aidianyehuangma" was selected as the target of GA_3 stress treatment. The results showed that the expressions of most genes were significantly increased or decreased, indicated that these CcWRKY transcription factor genes were likely to be involved in the defense response under GA_3 stress. At present, it has been shown that gibberellin-mediated signaling cascade regulates cellulose synthesis [48].Johnson et al. have reported that WRKY transcription factors were involved in epidermal hair development, glucose and gibberellin signaling. Therefore it will be very interesting to further studied about the relationships between CcWRKY genes, GA_3 stress and fiber development.

In addition, due to the importance of jute economy, it will be very important to study the transcriptional regulation of WRKY proteins in jute. These CcWRKY genes have positive effects on fiber development and response to GA_3 stress, and could promote fiber yield. At present, the research of transcription factors in the model plants was only a part of them, and it couldn't fully reveal the functions of each transcription factor. And the same transcription factors have different functions in different species because of the

different evolutionary directions. Overall, this study provided a theoretical basis for the subsequent study of CcWRKY family and its effects on fiber development by identification and bioinformatics analysis of CcWRKY gene family.

Conclusion

In this study, we have identified the CcWRKY gene family by analyzing gene structures, subgroup classifications, phylogenetic relationships and three-dimensional structures of protein. The CcWRKY proteins were highly conserved, the lengths of the gene sequences and the numbers of introns and exons varied widely, all WRKY genes showed a variety of expression patterns at different stages of stem development. In response to GA_3 stress, the fiber yield of a dwarf variety was increased. At the same time, the expressions of CcWRKYs have changed significantly like other fiber related marker genes. This suggested that CcWRKYs might be involved in the growth and development of bast fiber in jute.

Abbreviations

GA_3: Gibberellin.

Declarations

Acknowledgments

Not applicable.

Funding

This study was supported by the National Natural Science Foundation of China (31771369) and the National Agri-Industry Technology Research System for Crops of Bast and Leaf Fiber, China (nycytx-19-E06), http://119.253.58.231/.The funding numbers provided the financial support to the research programs, but didn't involve in work design, data collection, analysis and preparation of the manuscript.

Availability of data and materials

The datasets supporting the results of this publication are included within the article and its additional files.

Authors' contributions

LZ designed and directed the research; XW and YX analyzed the data; LZ wrote the manuscript with input from all authors; SN, JQ and LZ contributed new germplasm and techniques. All authors edited and agreed on the final manuscript.

Ethics declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Rushton P, Somssich I, Ringler P, J Shen Q: **WRKY transcription factors**. *Trends in Plant Science* 2010, **15**(5):247-258.
2. Yang P, Chen Z: **A family of dispersed repetitive DNA sequences in tobacco contain clusters of W-box elements recognized by pathogen-induced WRKY DNA-binding proteins**. *Plant Science* 2001, **161**(4):655-664.
3. Burley SK, Kamada K: **Transcription factor complexes**. *Current Opinion in Structural Biology* 2002, **12**(2):225-230.
4. Eulgem T, Rushton P, Robatzek S, Somssich I: **The WRKY superfamily of plant transcription factors**. *Trends Plant Sci* 2000, **5**(5):199-206.
5. Ling J, Jiang W, Zhang Y, Yu H, Mao Z, Gu X, Huang S, Xie B: **Genome-wide analysis of WRKY gene family in Cucumis sativus**. *BMC genomics* 2011, **12**(1):471.
6. Ishiguro S, Nakamura K: **Characterization of a cDNA encoding a novel DNA-binding protein, SPF1, that recognizes SP8 sequences in the 5' upstream regions of genes coding for sporamin and beta-amylase from sweet potato**. *Molecular and General Genetics MGG* 1994, **244**(6):563-571.
7. Riechmann JL, Heard J, Martin G, Reuber L, Z C, Jiang, Keddie J, Adam L, Pineda O, Ratcliffe OJ *et al*: **Arabidopsis Transcription Factors: Genome-Wide Comparative Analysis Among Eukaryotes**. *Science* 2000, **290**(5499):2105-2110.
8. Ramamoorthy R, Jiang S-Y, Kumar N, Venkatesh PN, Ramachandran S: **A comprehensive transcriptional profiling of the WRKY gene family in rice under various abiotic and phytohormone treatments**. *Plant & cell physiology* 2008, **49**(6):865-879.
9. Mangelsen E, Kilian J, Berendzen K, Kolukisaoglu Ü, Harter K, Jansson C, Wanke D: **Phylogenetic and comparative gene expression analysis of barley (Hordeum vulgare) WRKY transcription factor family reveals putatively retained functions between monocots and dicot**. *BMC genomics* 2008, **9**:194-200.
10. Yang B, Jiang Y-Q, Rahman M, Deyholos M, N V Kav N: **Identification and expression analysis of WRKY transcription factor in canola (Brassica napus L.) in response to fungal pathogens and hormone treatments**. *BMC plant biology* 2009, **9**(1):68-70.

11. Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelen JJ, Cheng J *et al*: **Genome sequence of the palaeopolyploid soybean**. *Nature* 2010, **465**(7278):178-183.
12. Wei K-F, Chen J, Chen Y-F, Wu L-J, Xie D-X: **Molecular Phylogenetic and Expression Analysis of the Complete WRKY Transcription Factor Family in Maize**. *DNA Research* 2012, **19**(2):153.
13. Dellagi A, Helibronn J Fau - Avrova AO, Avrova Ao Fau - Montesano M, Montesano M Fau - Palva ET, Palva Et Fau - Stewart HE, Stewart He Fau - Toth IK, Toth Ik Fau - Cooke DE, Cooke De Fau - Lyon GD, Lyon Gd Fau - Birch PR, Birch PR: **A potato gene encoding a WRKY-like transcription factor is induced in interactions with Erwinia carotovora subsp. atroseptica and Phytophthora infestans and is coregulated with class I endochitinase expression**. *Molecular plant-microbe interactions* 2000, **13**(10):1092-1101.
14. Wang Z, Yang P Fau - Fan B, Fan B Fau - Chen Z, Chen Z: **An oligo selection procedure for identification of sequence-specific DNA-binding activities associated with the plant defence response**. *The Plant Journal* 1998, **16**(4):515-522.
15. He H, Dong Q, Shao Y, Jiang H, Zhu S, Cheng B, Xiang Y: **Genome-wide survey and characterization of the WRKY gene family in Populus trichocarpa**. *Plant cell reports* 2012, **31**(7):1199-1217.
16. Dou L, Zhang X, Pang C, Song M, Wei H, Fan S, Yu S: **Genome-wide analysis of the WRKY gene family in cotton**. *Molecular Genetics and Genomics* 2014, **289**(6):1103-1121.
17. Wang L, Zhu W, Fang L, Sun X, Su L, Liang Z, Wang N, Londo JP, Li S, Xin H: **Genome-wide identification of WRKY family genes and their response to cold stress in Vitis vinifera**. *BMC plant biology* 2014, **14**(1):103.
18. Li J, Brader G, Palva ET: **The WRKY70 Transcription Factor: A Node of Convergence for Jasmonate-Mediated and Salicylate-Mediated Signals in Plant Defense**. *The Plant cell* 2004, **16**:319-331.
19. Rushton DL, Tripathi P, Rabara RC, Lin J, Ringler P, Boken AK, Langum TJ, Smidt L, Boomsma DD, Emme NJ *et al*: **WRKY transcription factors: key components in abscisic acid signalling**. *Plant biotechnology journal* 2012, **10**(1):2-11.
20. Zhang J, Peng Y, Sun Z-X: **Constitutive expression of pathogen-inducible OsWRKY31 enhances disease resistance and affects root growth and auxin response in transgenic rice plants**. *Cell research* 2008, **18**(4):508-521.
21. Devaiah BN, Karthikeyan AS, Raghothama KG: **WRKY75 transcription factor is a modulator of phosphate acquisition and root development in Arabidopsis**. *Plant physiology* 2007, **143**(4):1789-1801.
22. Xie Z, Zhang ZL, Zou X, Yang G, Komatsu S, Shen QJ: **Interactions of two abscisic-acid induced WRKY genes in repressing gibberellin signaling in aleurone cells**. *Plant J* 2006, **46**(2):231-242.
23. Islam MS, Saito JA, Emdad EM, Ahmed B, Islam MM, Halim A, Hossen QMM, Hossain MZ, Rasel Ahmed MSH, Kabir SMT *et al*: **Comparative genomics of two jute species and insight into fibre biogenesis**. *Nature plants* 2017, **3**(1):16223.
24. Grishanov SA, Harwood RJ, Booth I: **A method of estimating the single flax fibre fineness using data from the LaserScan system**. *Industrial Crops and Products* 2006, **23**(3):273-287.

25. Al-Ghazi Y, Bourot S, Arioli T, S Dennis E, J Llewellyn D: **Transcript Profiling During Fiber Development Identifies Pathways in Secondary Metabolism and Cell Wall Structure That May Contribute to Cotton Fiber Quality**. *Plant & cell physiology* 2009, **50**(7):1364.
26. Zhang L, Ming R, Zhang J, Tao A, Fang P, Qi J: **De novo transcriptome sequence and identification of major bast-related genes involved in cellulose biosynthesis in jute (Corchorus capsularis L.)**. *BMC genomics* 2015, **16**(1):164-179.
27. Tian Y, Dong Q, Ji Z, Chi F, Cong P, Zhou Z: **Genome-wide identification and analysis of the MADS-box gene family in apple**. *Gene* 2015, **555**(2):277.
28. Arnold K, Bordoli L, Kopp J, Schwede T: **The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling**. *Bioinformatics* 2006, **22**(2):195-201.
29. Kumar S, Stecher G, Tamura K: **MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets**. *Molecular Biology & Evolution* 2016, **33**(7):1872.
30. Green CE, Phillips RL: **Plant regeneration from tissue cultures of maize**. *Crop Science* 1975, **15**:417-421.
31. Niu X, Qi J, Zhang G, Xu J, Tao A, Fang P, Su J: **Selection of reliable reference genes for quantitative real-time PCR gene expression analysis in Jute (Corchorus capsularis) under stress treatments**. *Frontiers in plant science* 2015, **6**:848.
32. Livak K, Schmittgen T: **Analysis of Relative Gene Expression Data using Real-Time Quantitative PCR**. *Methods (San Diego, Calif)* 2002, **25**(4):402-408.
33. Duan MR NJ, LiangYH, Mao P, Lu L, Li LF, Wei CH, Lai LH, Li Y, Su XD: **DNA binding mechanism revealed by high resolution crystal structure of Arabidopsis thaliana WRKY1 protein**. *Nucleic acids research* 2007, **35**(4):1145–1154.
34. Wang H AU, Nakashima J, Hahn M G, Chen F, Dixon R A.: **Mutation of WRKY transcription factors initiates pith secondary wall formation and increases stem biomass in dicotyledonous plants**. *Proceedings of the National Academy of Sciences of the United States of America* 2010, **107**(51):22338-22343.
35. GUO Guang-yan BF, LIU Wei, BI Cai-li: **Advances in Research of the Regulation of Transcription Factors of Lignin Biosynthesis**. *Scientia Agricultura Sinica* 2015, **48**(7):1277-1287.
36. ZHANG Li-wu HW-x, TAO Ai-fen, LIN Li-hui, XU Jian-tang, FANG Ping-ping, QI Jian-min: **Research of Exogenous Hormone Sensitivity on Plant Height of a Dwarf Germplasm in Jute (Corchorus capsularis L.)**. *Scientia Agricultura Sinica* 2015, **48**(10):1892-1899.
37. Ulker B, Somssich I: **Ülker B, Somssich IE. WRKY transcription factors: from DNA binding towards biological function. Curr Opin Plant Biol 7: 491-498**. *Current opinion in plant biology* 2004, **7**(5):491-498.
38. Xie Z, Zhang Z-L, Zou X, Huang J, Ruas P, Thompson D, J Shen Q: **Annotations and Functional Analyses of the Rice WRKY Gene Superfamily Reveal Positive and Negative Regulators of Abscisic Acid Signaling in Aleurone Cells**. *Plant physiology* 2005, **137**(1):176-189.

39. Eulgem T, Somssich I: **Networks of WRKY transcription factors in defense signaling**. *Current opinion in plant biology* 2007, **10**(4):366-371.
40. KONG WeiLong YK, DAN NaiZhen, YANG ShaoZong, BAO ManZhu, HUANG XiangRong, FU XiaoPeng: **Genome-Wide Identification and Expression Analysis of WRKY Transcription Factor Under Abiotic Stress in Beta vulgaris**. *Scientia Agricultura Sinica* 2017, **50**(17):3259-3273.
41. Yang B, Jiang Y, Rahman MH, Deyholos MK, Kav NNV: **Identification and expression analysis of WRKY transcription factor genes in canola (Brassica napus L.) in response to fungal pathogens and hormone treatments**. *BMC plant biology* 2009, **9**(1):68.
42. Li H-L, Zhang L-B, Guo D, Li C-Z, Peng S-Q: **Identification and expression profiles of the WRKY transcription factor family in**. *Gene* 2012, **503**(2):248.
43. Yu Y, Wang N, Hu R, Xiang F: **Genome-wide identification of soybean WRKY transcription factors in response to salt stress**. *SpringerPlus* 2016, **5**(1):920.
44. Huang S, Gao Y, Liu J, Peng X, Niu X, Fei Z, Cao S, Liu Y: **Genome-wide analysis of WRKY transcription factors in Solanum lycopersicum**. *Molecular Genetics and Genomics* 2012, **287**(6):495-513.
45. Tang J, Wang F, Hou X-L, Wang Z, Huang Z-N: **Genome-Wide Fractionation and Identification of WRKY Transcription Factors in Chinese Cabbage (Brassica rapa ssp. pekinensis) Reveals Collinearity and Their Expression Patterns Under Abiotic and Biotic Stresses**. *Plant Molecular Biology Reporter* 2014, **32**(4):781-795.
46. Ling-ling D, Ya-ning G, Evans O, Chao-you P, Heng-ling W, Mei-zhen S, Shu-li F, Shu-xun Y: **Identification and expression analysis of group III WRKY transcription factors in cotton**. *Journal of Integrative Agriculture* 2016, **15**(11):2469-2480.
47. Wang H, Avci U, Nakashima J, Hahn MG, Chen F, Dixon RA: **Mutation of WRKY transcription factors initiates pith secondary wall formation and increases stem biomass in dicotyledonous plants**. *Proceedings of the National Academy of Sciences of the United States of America* 2010, **107**(51):22338-22343.
48. Huang D, Wang S, Zhang B, Shang-Guan K, Shi Y, Zhang D, Liu X, Wu K, Xu Z, Fu X *et al*: **A Gibberellin-Mediated DELLA-NAC Signaling Cascade Regulates Cellulose Synthesis in Rice**. *The Plant cell* 2015, **27**(6):1681-1696.

Tables

Due to technical limitations, tables are available as a download in the supplemental files section

Supplementary File Legends

Additional file 1: Table S1. Sampling method.

Additional file 2: Table S2. Variation of the heptapeptide WRKYGQK and zinc-finger structure in WRKY domains of CcWRKYs.

Additional file 3: Fig. S1. Protein 3D structure prediction of WRKY family in jute.

Additional file 4: Fig. S2. Expression profiles of CcWRKY genes in different tissues.

Additional file 5: Fig. S3. Comparison of average hypocotyl length of the seedling of "Huangma 179" and "Aidianyehuangma" treated with GA_3.

Additional file 6: Fig. S4. Expression profiles of CcWRKY genes under GA_3 stress.

Additional file 7: Fig. S5. Expression profiles of fiber marker genes under GA_3 stress (CesA: Cellulose synthase; CCoAOMT: Caffeioyl coenzyme A methyltransferase; 4CL: 4-Coumarate: Coenzyme A Ligase).

Additional file 8: Fig. S6. qRT-PCR verification of randomly selected 9 CcWRKY genes (Test varieties: "Huangma 179" and "Aidianyehuangma"; samples: jute stem bark of 60 days after sowing; CK: control, no treatment; 72h: 72h after GA_3 stress).

Figures

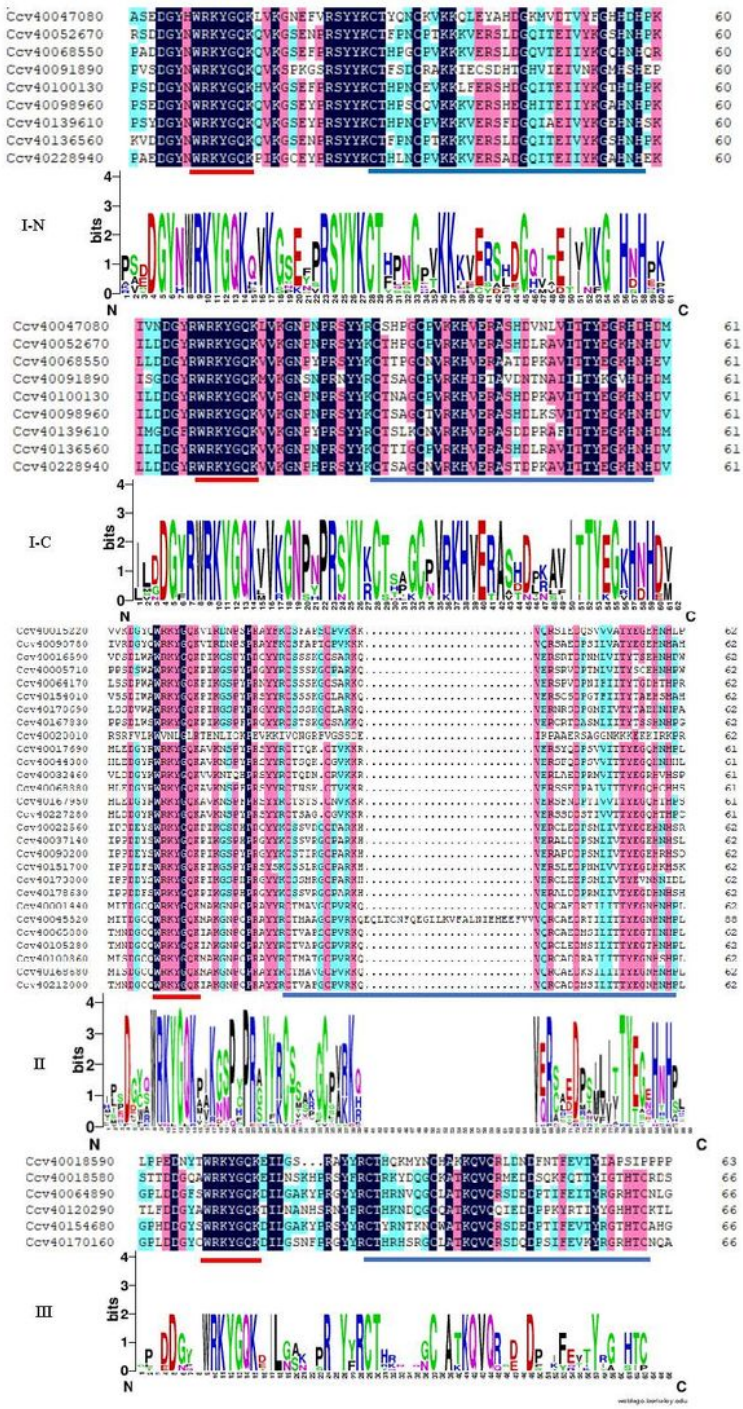

Figure 1

Sequence analysis of WRKY conserved domain of WRKY protein in jute

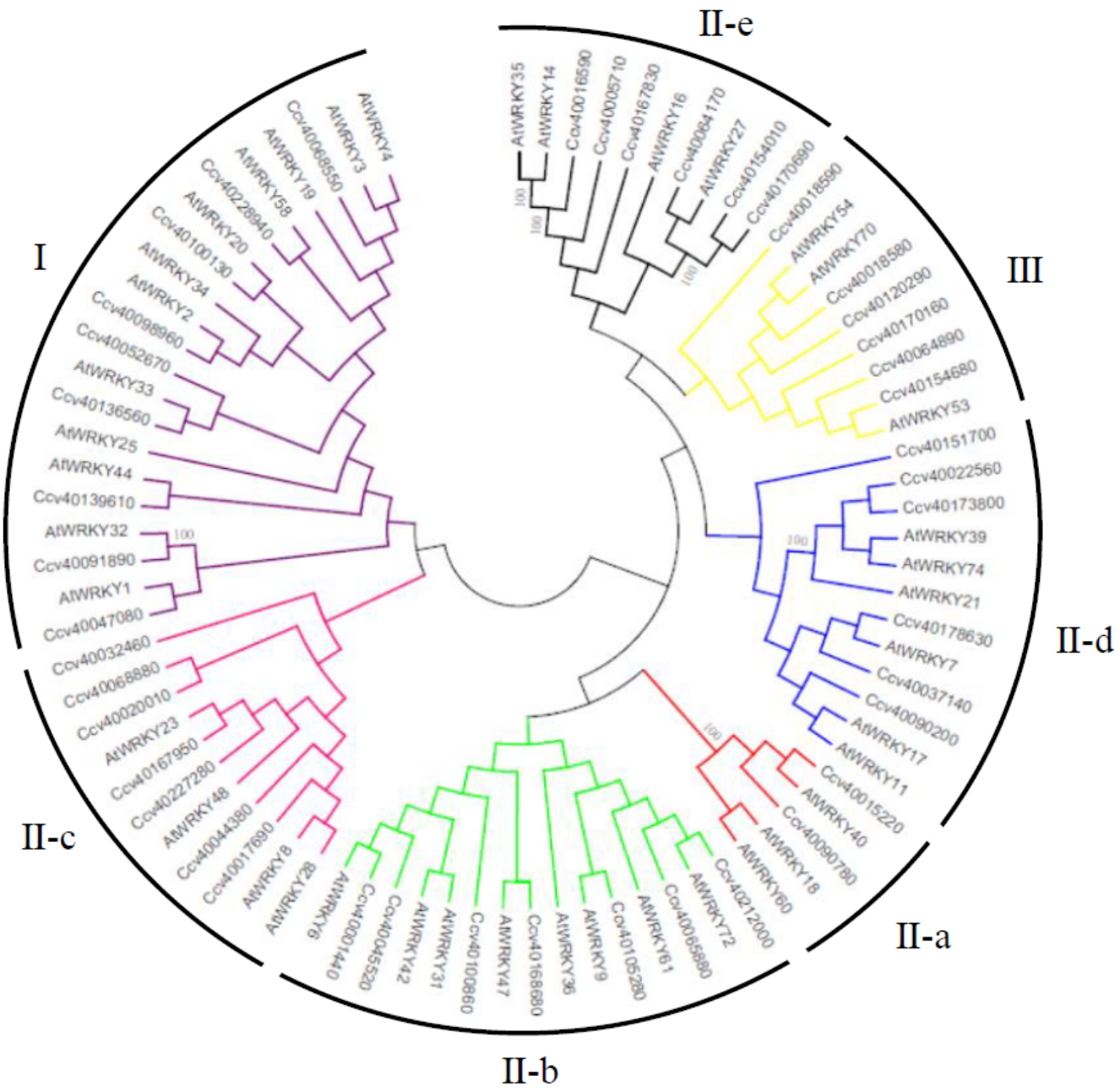

Figure 2

Phylogenetic tree of jute and Arabidopsis WRKY protein

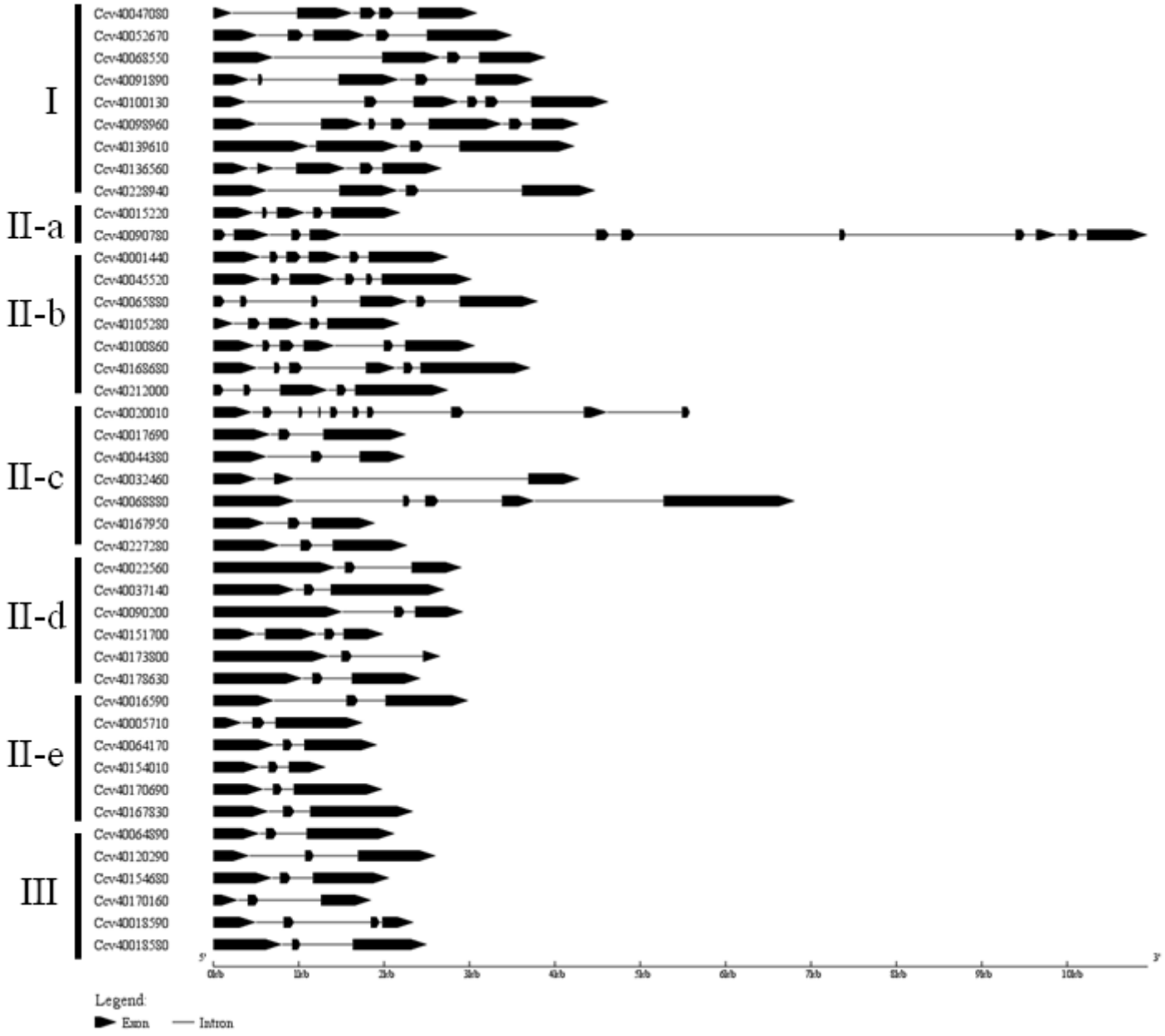

Figure 3

Intron and exon structure of CcWRKY genes in jute

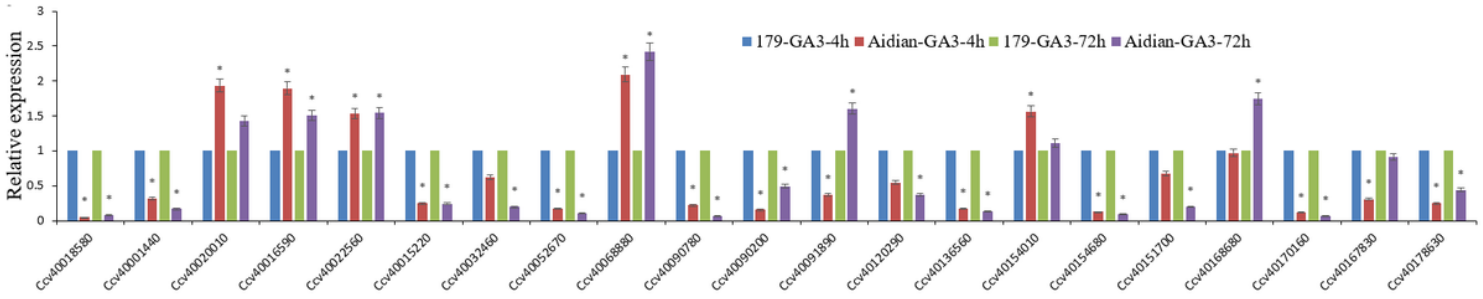

Figure 4

Effects of GA3 stress dwarf variety "Aidianyehuangma" on the expression of CcWRKY genes

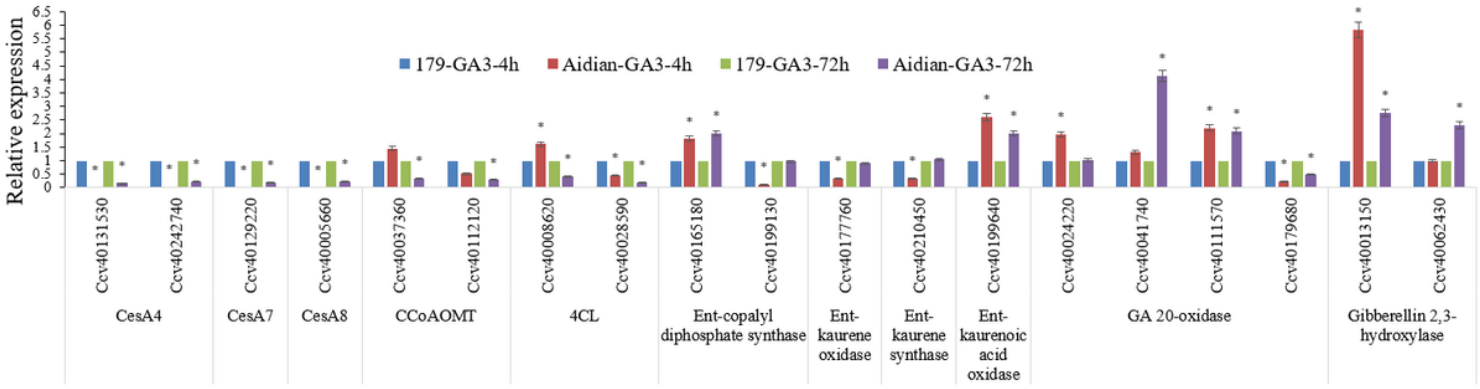

Figure 5

Effects of GA3 stress dwarf variety "Aidianyehuangma" on the expression of fiber related marker genes

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- FigS3.png
- FigS2.png
- TableS1.docx
- Table1.docx
- FigS1.png
- FigS6.png
- TableS2.docx
- Table2.docx
- FigS4.tif
- FigS5.tif