

Genome wide identification and functional characterization of the cotton C-repeat binding factor (CBF) under cold stress condition

Jiangna Liu

Zhengzhou Research Base, State Key Laboratory of Cotton Biology, Zhengzhou University/ Institute of Cotton Research, Chinese Academy of Agricultural Science, Anyang, China

Richard Magwanga

School of Biological, Physical Sciences, Mathematics and Actuarial Sciences (SBPMAS), Main Campus, Jaramogi Oginga Odinga University of Science and Technology (JOOUST), Main Campus, P.O. Box 210-4060

Yanchao Xu

Zhengzhou Research Base, State Key Laboratory of Cotton Biology, Zhengzhou University/ Institute of Cotton Research, Chinese Academy of Agricultural Science, Anyang, China

Tingting Wei

Zhengzhou Research Base, State Key Laboratory of Cotton Biology, Zhengzhou University/ Institute of Cotton Research, Chinese Academy of Agricultural Science, Anyang, China

Jie Zheng

Zhengzhou Research Base, State Key Laboratory of Cotton Biology, Zhengzhou University/ Institute of Cotton Research, Chinese Academy of Agricultural Science, Anyang, China

Yuqing Hou

Zhengzhou Research Base, State Key Laboratory of Cotton Biology, Zhengzhou University/ Institute of Cotton Research, Chinese Academy of Agricultural Science, Anyang, China

Yuhong Wang

Zhengzhou Research Base, State Key Laboratory of Cotton Biology, Zhengzhou University/ Institute of Cotton Research, Chinese Academy of Agricultural Science, Anyang, China

Erick Okuto

School of Biological, Physical Sciences, Mathematics and Actuarial Sciences (SBPMAS), Main Campus, Jaramogi Oginga Odinga University of Science and Technology (JOOUST), Main Campus, P.O. Box 210-4060

Stephen Gaya Agong

School of Biological, Physical Sciences, Mathematics and Actuarial Sciences (SBPMAS), Main Campus, Jaramogi Oginga Odinga University of Science and Technology (JOOUST), Main Campus, P.O. Box 210-4060

Kunbo Wang

Zhengzhou Research Base, State Key Laboratory of Cotton Biology, Zhengzhou University/ Institute of Cotton Research, Chinese Academy of Agricultural Science, Anyang, China

Zhongli Zhou

Zhengzhou Research Base, State Key Laboratory of Cotton Biology, Zhengzhou University/ Institute of Cotton Research, Chinese Academy of Agricultural Science, Anyang, China

Xiaoyan Cai

Zhengzhou Research Base, State Key Laboratory of Cotton Biology, Zhengzhou University/ Institute of Cotton Research, Chinese Academy of Agricultural Science, Anyang, China

Fang Liu (✉ liufcri@163.com)

Zhengzhou Research Base, State Key Laboratory of Cotton Biology, Zhengzhou University/ Institute of Cotton Research, Chinese Academy of Agricultural Science, Anyang, China

Research Article

Keywords: Cotton spp., QTL, Genetic map, Drought tolerance, miRNAs, Gene Ontology

Posted Date: April 23rd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-439424/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background Low temperature is a common biological abiotic stress in major cotton growing areas. Cold stress significantly affects the growth, yield and yield quality of cotton. Therefore, it is important to develop a more robust and cold stress tolerant cotton germplasm. In response to climate change and erratic conditions, plants have evolved various survival mechanisms, one of which induction of various stress responsive transcription factors, such as the C-repeat binding factors (CBFs), which have been found to enhance cold tolerance in various plants.

Results In this study detailed evaluation of the cotton CBF has been carried out. A total of 29, 28, 25, 21, 30, 26 and 15 proteins encoded by the CBF genes were identified in *Gossypium herbaceum*, *Gossypium arboreum*, *Gossypium thurberi*, *Gossypium raimondii*, *Gossypium turneri*, *Gossypium longicalyx* and *Gossypium australe*, respectively. Phylogenetic evaluation revealed that the proteins were grouped into seven clades, with clade 1 and 6 being the largest. Moreover, majority of the proteins encoded by the genes were predicted to be located within the nucleus, while some were distributed in other parts of the cell. Based on the transcriptome and RT-qPCR analysis, Gthu17439 (GthCBF4) was highly upregulated under cold stress, and was further validated through forward genetics. The Gthu17439 (GthCBF4) overexpressed plants showed a significantly tolerance to cold stress, with higher germination rate, higher root growth and high induction levels of stress responsive genes. The over-expressed plants exhibited low level of oxidative damage, due to significant reduction in the H₂O₂ production.

Conclusion The results showed that the Gthu17439 (GthCBF4) could be playing a significant role in enhancing cold stress tolerance in cotton and can be further exploited in developing cotton germplasm with an improved cold-stress tolerance

Background

Cotton is a thermophilic crop and is more sensitive to low temperatures [1]; China being the major cotton growing countries globally, the site specific regions within China, such as Xinjiang, is often affected by cold which significantly results in negative effects on plant growth and development [2]. Cold stress lead to inhibition of seed germination, reduction of plant growth and reproduction, as well as a decrease in crop yield and quality [3]. However, many crops, such as rice (*Oryza sativa*), maize (*Zea mays*), tomato (*Solanum lycopersicum*), soybean (*Glycine max*) and cotton (*Gossypium hirsutum*), do lack the ability to adapt to environments with low temperature, thus highly adaptive to tropical or subtropical regions [4–6]. In order to deter the emergence of the adverse effects, plants have evolved complex mechanisms to resist cold stress by integrating the plants transcription factors [7–10]. Through integration of the physiological, biochemical and molecular regulatory mechanisms to stress detection, signal transduction, gene expression and metabolic modifications, plants have been found to have survival strategy in reducing or preventing the effects of cell oxidative damage [11].

In order to understand the effect of cold stress on plants, it is important to distinguish between cold (0–15°C) and freezing/chilling stress (< 0°C). Cold stress leads to metabolic injury, destruction of the stability of protein complexes, inhibition of the metabolic pathways and various cellular processes in varying degrees, and eventually damage to the photosynthetic processes [12–14]. Freezing/chilling temperature promotes the ice formation in intercellular spaces of plant tissues, significantly reduces the efficiency of photosynthesis, and increased release of reactive oxygen species (ROS). Increased release of ROS leads to chloroplast membrane oxidation, moreover, increased levels of ROS initiate a stress signal mechanisms in plants which significantly alter gene expression at chloroplast and nucleus level thereby promoting lipid homeostasis, thereby increasing plants adaptation to chilling stress [15]. Furthermore, when ice crystals are deposited on the cell wall, the extracellular water potential decreases and the cell membrane is destroyed, resulting in serious cell dehydration [16]. The most common phenomenon to resist the adverse effects of cold is acclimatization through the evolution of stress responsive genes, for instance, the C-repeat binding factors (CBFs), being the CBF1 and CBF3 play a significant role in cold acclimation-dependent freezing tolerance [17]. Many plants have increased freezing resistance after a period of low non-freezing temperature, a phenomenon known as cold acclimation. Cold acclimation is an effective way to increase plant freezing resistance after a period of low non-freezing temperature [18]. When plants are exposed to chilling conditions, the plants do mobilize the cold-response genes (COR), which then activates.

C-repeat binding factors (CBFs), followed by the accumulation of cryoprotectants, which results in the acquisition of freezing tolerance [11]. When plants are exposed to nonfreezing cold condition, the CBFs are rapidly induced by low temperature, and then the downstream target gene *COR* become activated [19]. In addition, as the expression of CBF is regulated by light quality, biological clock and photoperiod, it is necessary to understand the daily and seasonal regulation of the *CBF* genes [20]. The C-repeat binding factor/dehydration-responsive element binding factor 1 (CBF/DREB1) is well-studied as an integral transcription factor in the cold regulatory pathway. It is an adaptive response where plants increase freezing tolerance after exposure to low non-freezing temperatures [21]. The CBF/DREB protein, a DNA binding protein belonging to AP2/ERF superfamily, was first identified in *Arabidopsis thaliana* as a cold (LT)-induced transcription factor [22]. APETALA2/ethylene-responsive element binding factor (AP2/ERF) family is a large class of plant-specific transcription factors, including AP2, RAV, ERF and CBF [23,24]. The expression of *CBF* gene was very low at room temperature, but increased rapidly within 15 minutes of plants exposure to cold stimulation [4,18]. During cold acclimation, mobilization of the CBF activated cold response (*COR*) gene and increased accumulation levels of the cryoprotectants, do significantly enhance the cold tolerance in plants [11]. The *CBF / DREB* gene is activated by *CBF* expression (ICE) inducer through specific binding of *cis*-elements to MYC (c-Myc, L-Myc, S-Myc, and N-Myc) in the promoter region [18,22,25].

The cold reaction pathway of ICE-CBF-COR has been proved to be the effective defense mechanism to cold stress [4]. Moreover, the known six members of the CBF family members CBF1/DREB1C, CBF2/DREB1B, and CBF3/DREB1A are induced by cold stress, while CBF4/DREB1D, DREB1E/DDF2, and

DREB1F/DDF1 are induced by osmotic stresses [26]. Furthermore, the CBF1, CBF2 and CBF3 are arranged in tandem in an 8.7-kb region of the short arm of chromosome 4 in Arabidopsis [22]. All the three CBF Transcription factors perhaps acts as the activator of the expression of downstream *CORs* genes, but with different functions [27]. The CBF1/2/3 triple mutant Arabidopsis seedlings obtained by CRISPR/Cas9 technology are more sensitive to low temperature than the single mutant CBF2 and CBF3, as well as the double mutant CBF1/3 [17], which is an indication that *CBF* genes play significant role enhancing cold stress tolerance in plants. .

The CBF Transcription factors have been widely identified and isolated from rice, tomato, *Brassica napus*, wheat, barley and maize, which show that the CBF family is large in scale and complex in structure [2,28]. It has been reported that *AtCBF* gene was overexpressed in *Brassica napus* transgenic plants, and improved the cold tolerance [2,29]. The combination of *AtDREB1A* gene with stress-induced *RD29A* promoter improved the tolerance in transgenic tobacco to drought and cold stress [30,31]. Through phylogenetic analysis of Arabidopsis CBFs and its orthologous genes in other plants, it was found that CBFs were highly conservative in phylogeny [18]. In understanding cold stress response among the cotton genotypes, Cai et al (2019) found that the cotton germplasms of the D genome respond differently to drought stress, in which *G. thurberi* exhibit higher levels of cold stress tolerance by mobilizing the antioxidant enzymes and induction of stress responsive genes such as the *CBF4* and *ICE2*. As an important oil and fiber crop, cotton has been planted in more than 70 countries and plays an important role in the global economy. However, cotton yield is often adversely affected by abiotic stresses. Therefore, studying the molecular adaptation mechanism to stress resistance in cotton is of great significance for improving cotton yield. The 21 *CBF* genes have been cloned from *G. hirsutum*, which provides useful clues for understanding the cold tolerance mechanism in cotton. However, due to the limited genome sequence, the expression profile of *CBF* family and its phylogenetic relationship with other plant the role of the *CBF* members are still unclear. In order to better understand the function and evolutionary relationship of the *CBF* gene family in cotton, we analyzed the structural variation and evolution pattern of *CBF* family based on the genome-wide data of several cotton species, and explored the molecular mechanism of cold adaptation formation in *G. thurberi*. This study provides fundamental information and reference for further research on the molecular mechanism of the *CBF* genes and their regulatory role to cold adaptation in cotton.

Results

Identification of CBF family genes in the cotton genome

The availability of the whole sequences for the seven cotton species enabled us to identify the CBF proteins harbored in their genome. The Pfam domain PF00847 was used as the query to obtain the CBF proteins, and finally 29 members of the *CBF* genes were identified in *G. herbaceum*, 28 members in *G. arboreum*, 25 members in *G. thurberi*, 21 members in *G. raimondii*, 30 members in *G. turneri*, *G. longicalyx* had 26 members while *G. australe* had 15 *CBF* genes. Three representative cotton species from these seven species were then chosen for further detailed analysis: *G. herbaceum*, *G. thurberi* and *G. australe*.

The CBF CDS length in *G. herbaceum* ranged from 306 bp to 1,230 bp, while for *G. australe*, CDS ranged from 429 bp to 1,077 bp. In the analysis of the physiochemical properties of the CBF proteins, the results show a great difference, for instance, the CBF proteins obtained from the *G. herbaceum*, their molecular weights (MW) ranged from 11,241.88 Da to 39,020.73 Da; isoelectric value (*pI*) ranged from 5.26 to 10.27. While in *G. thurberi*, the MW of the proteins encoded by the *CBF* genes ranged from 16.17923kDa to 45.59153kDa, the *pI* ranged from 5.11 to 10.83. Similarly in *G. australe*, it ranged from 15.67636kDa to 40.86796kDa. The *CBF* genes differed substantially by the encoded protein size and its biophysical properties (Table S1).

Phylogenetic analysis of cotton *CBF* gene family

In order to determine the phylogenetic relationship of the CBF proteins, we constructed a phylogenetic tree by MEGA7.0, using the Neighbor-joining (NJ) method with minimal evolution and maximum parsimony. The CBF proteins were clustered into 7 clades and designated as clade 1 to 7 (Fig. 1). Clade 1 contained 61 CBF protein sequences at most, while clade 3 contains only 7 CBF amino acid sequences. Consistent with previous classification, all of the Arabidopsis CBFs were distributed among the clade 6 [2,24,29]. Except *G. australe* did not appear in clade 3, the other cotton species were distributed in all 7 groups.

Chromosomal mapping, Gene structure and C-terminal conserved motifs analysis

All the genes located on various chromosomes in the three cotton genomes, were named according to their position on the chromosome. In *G. herbaceum* a member of the diploid type of the A genome, the *CBF* genes were mapped on all the chromosomes, except chromosome Chr01 which harbored no *CBF* genes. The highest gene loci were observed in chromosome Chr05 and Chr07 with 5 and 6 genes respectively, while the lowest gene loci was observed in chromosome Chr04,06,08 and Chro09 with a single gene locus in each (Figure 2A). In the diploid of the D genome, *G. thurberi*, chromosome Gthu_1, Gthu_8 and Gthu_9 harbored no genes, while chromosomes, Gthu_5, Gthu_7, and Gthu_12 had more gene loci, however, the highest gene loci was noted in chromosome Gthu_05, 07 and 12 with 4, 6 and 5 *CBF* genes respectively, similarly lowest gene loci was noted on chromosome Gthu_3, 6 and 13 with a single gene locus each (Fig 2B). Finally, in *G. australe* diploid cotton of the G genome, no *CBF* genes were found in chromosome G6, G9, G11 and G12, but the highest gene loci was only observed in chromosome G7 with 5 genes (Figure 2C).

In order to analyze the motifs, we employed MEME to detect conserved motifs in the *CBF* family. There were 10 conserved motifs distributed in each CBF family (Figure 3A). Almost all CBF proteins had the same number of motifs; three motifs were prominent among the CBF proteins, motifs 1, 2, and 4. Analyzing the arrangement of exons and introns can provide important insights into the evolution of gene families [13]. To study the exon/intron structure of the *CBF* gene, the CDS and the genome sequence were compared. The results showed that most of *G. herbaceum* and *G. australe* harbored no intron, and indication that the genes were highly conserved, but for the *G. thurberi* the least number of exons were two but others contained as higher number (Figure 3B-E). Basically members of the same group phylogenetically harbored similar number of intron-exon ratio. Stress regulates the expression of primary

and specialized metabolism genes at the transcriptional level via transcription factors binding to specific cis-elements, a number of *cis*-regulatory elements were found to be associated with the various cotton *CBF* genes, for instance MYB, ABA-responsive element (ABRE), long term repeat (LTR) among others (Figure 3F), the MYB, ABRE are among the top ranked stress responsive *cis*-regulatory elements [33]. The detection of myriad of *cis*-regulatory elements reveals the significant role played by the members of the cotton *CBF* genes in enhancing stress tolerance, more so cold stress in plants.

In determining the possible cellular sublocalization of the proteins encoded by *CBF* genes, an online tool wolf sport was employed. Among the three cotton species, the highest proportions of CBF proteins were embedded in the nucleus. Moreover, the nucleus is the central regulator of the cellular activities. However, other signals were observed in other cellular compartments, such as the chloroplast, mitochondria, plasma membrane, vacuole membrane and chloroplast (Table S2)

RNA-seq analysis and RT-qPCR validation of the *CBF* genes under cold stress conditions *G. thurberi* transcriptome data was used to analyze the expression patterns of the 24 *CBF* genes under cold stress (Figure 4A). Among them, only 17 genes were differentially expressed under cold stress. According to *G. thurberi* transcriptome data (previous work by the research team, yet to be published), 12 differentially expressed *CBF* genes were selected, and 15 and 13 genes were selected from the two cotton varieties *G. herbaceum* and *G. australe* based on the RNA sequence data. RT-qPCR was employed to analyze the expression pattern; most genes were up-regulated in the three cotton species (Figure 4B). Among the 12 genes in *G. thurberi*, 10 genes exhibited significantly higher upregulation, with only 2 genes being down-regulated. The expression trend was consistent with the transcriptome data. In *G. herbaceum*, 9 genes were up-regulated, and 6 were down-regulated. In *G. australe*, 8 were up-regulated and 5 were down-regulated. The higher level of upregulation among the *CBF* genes across the various cotton types is an indicator that the proteins encoded by the *CBF* genes could be playing an integral role in enhancing cold acclimatization in cotton plants. Moreover, by integrating the transcriptome data and RT-qPCR result; we selected one of the highly up regulated *CBF* genes, *GthCBF12.5* (*GthCBF4*) for further validation.

Experimental validation of subcellular localization of cotton CBF proteins

The results showed that the none pCAMBIA2300-eGFP-Flag-*GthCBF4* infused plants showed green fluorescence signals on both the nucleus and cell membrane, while the fusion protein pCAMBIA2300-eGFP-Flag-*GthCBF4* only had green fluorescence signals in the nucleus (Figure 5), indicating that the protein encoded by the gene was localized in the nucleus. The results were in agreement to the bioinformatics prediction of the possible cell compartmentalization of the CBF proteins.

Phenotype and cell damage identification of *GthCBF4* overexpressed Arabidopsis under low temperature.

Two highly expression lines OE-1 and OE-3 were selected for phenotypic evaluation (Figure 6A), phenotypically, the *GthCBF4* overexpressed lines (OE-1 and OE-3) showed no significant difference with the wild types under normal conditions, but when the plants were exposed to cold stress, the survival of the WT were significantly reduced, while the OE-1 and OE-3 showed higher level of survival (Figure 6B), in

which the OE plants survival rate was estimated at 60% compared to the WT with only 2% survival rate (Figure 6C). In determining the expression levels of the *GthCBF4* overexpressed gene, at 0h, the WT showed no expression, but at 1h and 3h of cold stress exposure, the *CBF* genes were partially induced but insignificantly, slightly below one, while the reverse was observed among the OE plants. The expression levels of the *GthCBF4* overexpressed gene was significantly high, with 0h showing expression levels close to three folds, while at 1h and 3h, the expression levels of the *GthCBF4* overexpressed gene was above four folds (Figure 6D). Under normal conditions, the stained blue area on the leaves of the transgenic overexpression plants and wild-type plants was very small. While under cold treatment, the blue areas on transgenic leaves were significantly smaller than the wild-type, the color depth was also lighter. Furthermore, the DAB staining method was used to reflect the accumulation of H_2O_2 in Arabidopsis leaves. The accumulation of H_2O_2 in the transgenic overexpression leaves and the wild-type were very low under normal conditions, and the production of brown matter was hardly seen. But after cold treatment, the brown area on the wild type Arabidopsis leaves was obviously larger than the wild type, and the color depth was also deeper (Figure 6E). The results showed that the overexpression of the *CBF* gene in the transgenic lines improved the ability of the plants to oxidize the oxidative agents such as the H_2O_2 thereby reducing the levels of oxidative damage in the plants.

Evaluation of physio-morphological traits in *GthCBF4* Overexpressed and wild type plants under low temperature environment

Germination of the OE and WT lines showed no significant differences under normal or controlled conditions, however, under cold stress, none of the WT germinated while the OE lines showed some level of germination (Figure 7A-B), an indication that the OE lines were significantly improved and were able to adapt to cold stress condition. In the evaluation of the root lengths, no significant differences was observed under controlled conditions, but under cold stress conditions, the OE lines showed higher root lengths compared to the WT (Figure 7C-D), thus the overexpression of the *CBF* genes could be playing a role either in the rate of cell division and or cell elongation. We further evaluated known stress responsive genes such as the *COR15A*, *RD29A*, *KIN1* and *COR47* [17]. The OE lines significantly showed higher expression levels to all the stress responsive genes profiled (Figure 7E). The high induction levels of the stress responsive genes in the OE lines, showed that the overexpressed genes do not suppress the expression of the stress responsive genes but do promote their expression, an indication that the overexpressed gene could be playing a vital role in enhancing cold stress tolerance in plants.

Discussion

Cotton is an important economic crop, which supports the economies of a number of countries globally, moreover, it is the main source of natural fiber critical raw material for the textile industries [34], however, in the recent past, its global production has significantly shrunk due to climate change. Rainfall has become erratic and scarce, and even the average daily temperature has significantly increased with prolonged periods of cold weather, thereby affecting the growth and development of the plant. In order to improve the performance of current elite cultivars, a number of strategies have been proposed, among

them is the utilization of the plants transcription factors and other novel genes to improve the adaptability of the current cotton germplasms to ever changing environmental conditions. Cold stress has never been a major concern in cotton production, but the climate dynamics, has increased deleterious effects of cold stress, for instance, in the 1980s, united states of America loss due to chilling effects was estimated at 60 million dollars [35]. Furthermore, in very sensitive crops, cold stress causes considerable agricultural yield loss, particularly in sensitive crops like maize, rice and chickpea [36]. To increase the chance of survival, plants have adopted numerous mechanisms to cope with the ever changing environmental conditions, one of which is evolution of novel genes with tremendous effects in improving plants adaptability to various environmental stress factors, for instance the *LEA* genes have been found to offer drought stress tolerance in cotton [37].

Although abiotic stress is a major challenge in cotton growth, there is no detailed study on the *CBF* gene in cotton [38]. In previous studies, *CBF* family have been identified in cotton (*Gossypium hirsutum*) [26], wheat [39], lettuce [40], *Brassica napus* [41], Barley [42] and soybean [43], but there is few study in diploid cotton. In this work, genome wide identification, characterization and functional analysis of the proteins encoded by the cotton *CBF* genes was done, in which 29, 25 and 15 *CBF* proteins encoded by the *CBF* genes were identified in *G. herbaceum*, *G. thurberi* and *G. australe*, respectively. The results obtained showed that the proteins encoded by the *CBF* genes in cotton were higher compared to other plants such as lettuce with 14, *Brassica napus* with 10 and soybean with 14 *CBF* genes, but less than hexaploid wheat with 65 *CBF* genes and barley with 20 *CBF* genes. The high number of the proteins encoded by the *CBF* genes in cotton perhaps could explain the significant role they play in plants in relation to stress tolerance. Moreover, studies have found that Arabidopsis *CBF* family members have an AP2 domain, and each has a conserved amino acid sequence upstream and downstream of AP2. The upstream of AP2 is PKK/PKKPAGR (RAGRxxKFxETRHP) and the downstream is DSAWR [44]. If the PKK/PKKPAGR mutation can inhibit the binding ability of *CBF* to the *COR* promoter of its downstream genes, thereby weakening the level of *CBF* regulation, this sequence is necessary for *CBF* to perform its transcription factor function. Cotton genome contains a large and complex *CBF* subfamily, with a conserved AP2/EREBP domains and with *CBF*-like characteristics, indicating that cotton *CBFs* have similar function with other *CBFs* in dicotyledons [27]. Phylogenetic analysis showed that the *CBF* families are divided into seven groups, among these genes *CBF1*, *CBF2* and *CBF3* are all induced by cold stress in *Arabidopsis thaliana*. Therefore, we speculate that *CBF* genes may also respond to abiotic stress, especially to cold. Further analysis showed that the isolated *CBF* gene was highly expressed under cold stress, consistent with previous research results.

The cotton *CBF* genes phylogenetic analysis revealed that the entire members of the cotton *CBF* family were grouped into seven clades, this deviates from the previous findings in which five clusters were observed among the *CBF* members obtained from lettuce [40], but were in agreement to findings obtained for the Tea Plant (*Camellia sinensis*) in which the *CBF* and their homologs were classified into six groups [45]. In plants, the transcriptional regulation of osmotic stress response mainly depends on two main *cis*-regulatory elements, which are related to stress response genes ABREs and dehydration response elements (DREs). The two *cis*-regulatory elements were found to be associated with the *CBF* genes, which

is an indicator that the members of the *CBF* gene family could perhaps be playing a significant role in enhancing cold stress tolerance in plants. Moreover, the DREs are mainly involved in ABA-independent pathways [46], and ABRE is responsible for detecting ABA-mediated osmotic stress signals [47]. In this study, the *CBF* genes detected for the three cotton species were rich in stress response genes (ABRE), dehydration response element (DRE) and cold stress response element (LTR). It is suggested that exogenous environmental stress can induce the expression of *CBF* gene through its response to *cis*-regulatory elements, and further improve the resistance of plants to environmental stress.

It is universally accepted that transcription factors must be present in the nucleus to perform their functions [48]. In the evaluation of the possible cell compartmentalization of the proteins encoded by the *CBF* genes, both prediction and experimental evaluation showed that majority of the CBF proteins are localized within the nucleus. Several studies have found that a number of stress responsive proteins are sub-localised within the nucleus, for instance calmodulin-regulated proteins, in plants, the activity of a pea apyrase, PsNTP9, is localized in the nucleus, and have a significant role in enhancing resistance to heavy metals and xenobiotic Compounds [49]. Moreover, when plants exposed to environmental stress, there is increased release of oxidizing agents, which destroys the DNA, to counteract the DNA damage, plant cells evolved mechanisms for the DNA repair in both nucleus and mitochondria [50], thus the abundance of the CBF proteins in the nucleus could be significant in the DNA repair as a result of oxidative damage..

Overexpression of Arabidopsis *CBF* gene in other plant species or overexpression of *CBF* of other species in Arabidopsis has revealed the potential of the *CBF* genes in enhancing frost resistance [51]. Moreover, it has been shown that the *CBF* gene plays an important role in plant cold acclimatization, downregulation of *CBF1* and *CBF3* results in 25-50% reduction in tolerance to cold stress levels in pre-cold treated mutant plants [52]. In this study, the *GthCBF4* was strongly up-regulated in cotton seedlings under cold treatment. The survival rate of *GthCBF4* transgenic *Arabidopsis thaliana* plants was significantly improved after freezing treatment. Furthermore, the inducer of CBF Expression (ICE) is a pioneer of cold acclimation, an MYC-type basic helix-loop-helix family transcription factor (TF) [53]. It has been reported that ICE1 plays a significant role in C-repeat binding factor 3 (CBF3) cold induction [25]. When plants encounter cold stress, ICE1 could be mobilized from JAZs bound by DELLAs and trigger the expression of CBF3. The CBF3 activates the expression of GA2ox7 to reduce the bioactive gibberalic acid (GA) level, which enhance the accumulation of DELLAs. Therefore, DELLAs can control the cold induction of CBF3 through ICE1 via JAZs. Moreover, from the trypan blue staining and DAB staining revealed that the *GthCBF4* overexpressed plants had a significant reduction in cold injury compared to the wild types, a strong indicator of the possible roles of *CBF* genes in enhancing cold stress tolerance in plants. Moreover, root growth was significantly higher in *GthCBF4* overexpressed plants.

In order to further elucidate the critical roles played by the proteins encoded by the *CBF* genes, known stress responsive genes were profiled, and the results showed that overexpression of the *CBF* gene significantly increased the induction levels of the stress responsive genes. In general, the expression level of the four *COR* genes were significantly correlated with the freezing tolerance level. Furthermore,

COR15A, which encodes the chloroplast-targeted polypeptide, enhance the cold resistance of the chloroplasts [54]. In addition, the *CBF* genes have the ability to induce the expression of the *COR47*, *RD29A* and *KIN1* genes thereby improving the plants tolerance level to cold stress[17]. The up-regulation of the stress responsive genes further confirmed that the GthCBF4 overexpressed plants ability to tolerate the effects of cold was significantly higher compared to the wild types. In the past two decades, the transcriptional network of CBF signaling pathway have been extensively studied, and many *COR* genes have been identified in genome-wide expression profiles, and about 10-25% of them are regulated by CBF, which implies that more early cold regulated transcription factors are mainly involved in improving cold tolerance.

Conclusions

In conclusion, based on the genome wide identification and functional analysis of the proteins encoded by the CBF genes, a total of 69 CBF genes were identified in three diploid cotton species, in which 29, 25 and 15 CBF proteins in *G. herbaceum*, *G. thurberi* and *G. australe*, respectively. The RNA data and the expression profiling of the CBF genes revealed that higher proportions were upregulated under cold stress conditions. moreover, overexpression of the Gthu17439 (GthCBF4) gene in model plant, *Arabidopsis*, revealed that the GthCBF4 overexpressed plants ability to tolerate cold stress was higher compared to the wild types, as evident on the germination, root growth and induction of various stress responsive genes. The results therefore provides fundamental information for future exploration of the CBF genes in the development of a more robust cotton germplasms, which are highly resilient to cold and or chilling stress.

Materials And Methods

Identification of CBF family genes in the cotton genome

Wild diploid cotton species genome data was retrieved from the CottonGen (<https://www.cottongen.org/>) to construct a local BLAST database (<https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins>) . The *Arabidopsis* CBF protein sequences were used as probes to compare with the wild diploid cotton. The E-value threshold for BLASTP was set at $1e^{-10}$ to obtain the final dataset of the CBF proteins. The protein family Pfam (Finn et al., 2014), and a Simple Modular Architecture Research Tool (SMART) (<http://smart.embl-heidelberg.de/>) databases were employed to confirm each predicted CBF protein sequences [55]. Redundant sequences and incomplete sequences were removed. The sequences of 10 *Gossypioides kirkii* and 9 *Theobroma cacao* CBF proteins were obtained from the CottonGen (<https://www.cottongen.org/>) and national center for biotechnology information (NCBI) (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) databases, respectively. In addition, physicochemical parameters including the molecular weight (MW) and isoelectric point (*pI*) of each gene product were calculated using compute the pI/Mw tool from ExPASy (<http://www.expasy.org/tools/>), as previously used by Magwanga et al [37,56] in the evaluation of the physicochemical properties of the Late embryogenesis abundant (*LEA*) genes in cotton.

Sequence alignment and phylogenetic analysis of the cotton *CBF* gene family

An alignment of multiple CBF protein sequences from *A. thaliana*, *Gossypioides kirkii* and *Theobroma cacao* were generated using the ClustalW program (<https://www.genome.jp/tools-bin/clustalw>). A neighbor-joining (NJ) analysis of the generated alignment was performed using the unweighted pair-group method with arithmetic mean algorithm to construct an unrooted phylogenetic tree. Bootstrap value was 1000, and other parameters were used by default value. The tree was visualized with MEGA 7.0 software [57].

Gene structure and C-terminal conserved motifs analysis

Structural information for the *CBF* genes including the chromosomal location and gene length were determined. The exons and introns were predicted by comparing the coding sequences with genomic sequences. The conserved motif analysis of the CBF protein sequences were predicted by MEME online software (<https://meme-suite.org/meme/>). Moreover the Conserved Domains Database (CDD) (<https://www.ncbi.nlm.nih.gov/cdd/>) was employed to search for the conserved domain information of the CBF, and used the TBtools mapping tool to draw the conserved domains [58].

Retrieval and analysis of promoter sequences

The 2000 bp sequence upstream of ATG were extracted from the transcription start site of the *CBF* gene sequence, and submitted the obtained sequence to the PlantCARE website (<http://www.dna.affrc.go.jp/PLACE/signalscan.html>). Identification of possible cis-acting elements in the promoter region is used to identify putative cis-regulatory elements in the promoter sequence [59]. In addition, we carried out the subcellular localization prediction of all the CBF proteins by an online tool WoLFPSORT (<http://www.genscript.com/wolf-psort.html>) [60].

RNA extraction and qRT-PCR analysis

At three leaf stages, cold stress was imposed by subjected to cold stress by transferring the seedlings and kept at 4°C under normal light condition. The leaves were then harvested for RNA extraction at 0h, 0.5h, 3h, 6h, 12h and 24 h of post stress exposure. The total RNA was extracted using EASYspin plus plant RNA kit (Aidlab, Biotech, Beijing, China), following the manufacturer's instructions. The quality and concentration of each RNA sample was determined by Nano Drop 2000 spectrophotometer and RNA that fulfilled standard at 260/280 in a range of 1.80 – 2.1 was used for further analysis [38]. The primers used for qRT-PCR were designed using primer premier 5 software for all genes (Table S3). The cotton *GhActin* gene, forward primer sequence 5'ATCCTCCGTCTTGACCTTG3' and reverse primer sequence 5'TGTCCGTCAGGCAACTCAT3'), was used as a reference gene for the analysis. Real-time PCR reactions were carried out in a final volume of 25 µl, using a SYBR Green master mix and an ABI7500 thermal cycler (Applied Biosystems, Foster City, CA, U.S.A.), following manufacturer's instructions. The single stranded cDNA was synthesized from reverse transcriptase with TranScript-All-in-One First-Strand cDNA Synthesis SuperMix (TransGen Biotech kit Beijing, China) for RT-qPCR in harmony with manufacturer protocol. NCBI

primer blast was used to design 27 LHC gene primers. The primers details are given in (Table S1). The 7500 fast real time system was used for analysis and 10 μ L of Time fastStart Universal SYBR Green Master sigma (ROX) solution (Roche Diagnostics Mannheim, Germany) per sample was used for RT-qPCR. Total reaction mixture volume was 20 μ L, containing 2 μ L cDNA, 6 μ L RNA free water, 10 μ L green SYBER, 1 μ L of each forward and reverse primer. Ghactin7 was used as control [61]. The RT-qPCR completed with three technical repeats. The expression level of genes was calculated using the formula $E = 2^{-\Delta\Delta Ct}$.

Plant material

The seeds of *G. herbaceum*, *G. thurberi* and *G. australe* were pre-germinated in sand at 25°C for 4 days. Then seedlings were then transferred to the hydroponic facility equipped with Hoagland nutrient solution [62]. The greenhouse conditions were set at 28°C during the day/25°C at night, the photoperiod was 16 hours, and the relative humidity was 60-70%. At three-leaf stage, cotton seedlings were subjected to cold stress by transferring the seedlings and kept at 4°C under normal light condition. The leaves were then harvested for RNA extraction at 0h, 0.5h, 3h, 6h, 12h and 24 h of post stress exposure. Each treatment was repeated three times. The leaf samples were put in liquid nitrogen, frozen and stored at -80°C before RNA extraction.

GthCBF4 subcellular localization analysis

The *Agrobacterium tumefaciens* (LBA4404) strain *GthCBF4* vector was used for transformation studies. The design of GFP fusion constructs for pCAMBIA2300-eGFP (control GFP). The total RNA (1 μ g) isolated from leaves of *G. thurberi* was used for amplification of first strand of cDNA (TransGen Biotech Beijing). Specific primers were used to amplify, C-repeat binding factors protein (*GthCBF4*) gene products from *G. thurberi*. The oligonucleotides used for amplification of *GthCBF4* were forward primer 5'GAGAACACGGGGGACTCTAGAATGGTTGATTCTGGGTCGGTTTCT3' and reverse primer 5'ACCCATGTTAATTAAGGATCCAATAGAATAACTCCATAAAGG3' (sigma, Henan, China). The underlined sequences (TCTAGA and GGATCC) in the primer pair represent restriction sites for *Xba*I and *Bam*HI respectively. The reverse primers used for amplification of *GthCBF4* genes were designed in a fashion so as to eliminate stop codons of these genes. The PCR amplified products of *GthCBF4* were digested and cloned at the restriction site(s), *Xba*I and *Bam*HI respectively in the vector (GFP) backbone. The PCR cycling conditions were as follows: Step (1) 95°C-3 min, (2) 95°C-15 sec, (3) 58°C-15 sec, (4) 72°C-1 min, (5) cycle to step 2 to 4 for 35 more times, (6) 72°C-5 min, (7) incubate at 4°C. The PCR was performed using Phanta Max Super-Fidelity DNA polymerase (Vazyme, Nanjing, China). To explore the subcellular localization of the *GthCBF4* gene, a pCAMBIA2300-eGFP-Flag-*GthCBF4* fusion vector of CBF and GFP was transiently infused in the epidermal cells of tobacco leaves, the construct was driven by the 35s promoter and transformed *Agrobacterium* LBA4404 competent cells. In addition, *Agrobacterium* competent cells expressing only the *GFP* gene were used as a negative control. Four-week-old tobaccos cotyledon flat leaves were selected for infusion, and cultivated in dark for 24h-36h after infiltration, and then fluorescence observation was performed under a laser confocal microscope.

Cloning of *GthCBF4*, a CBF novel gene in arabidopsis

To explore the function of *GthCBF4* gene in plants. Constructed a pBI121-*GthCBF4* recombinant vector, transformed into the promoter cells of *A. tumefaciens* GV3101 [63] as previously applied by Magwanga et al [7] by adopting a freeze-thaw method [64] was performed. The wild-type (WT) *A. thaliana* were transformed by adopting the dipping method [65] by using a 50 mg/mL kanamycin for positive selection up to the third generation (T3) At the second generation, also referred to as the T2 generation, the *GthCBF4* overexpressed lines were identified, through polymerase chain reaction, two high-overexpressed transgenic lines were obtained, and grown to generate the stable T3 generation.

Phenotypic identification, physiological and biochemical parameters of WT and *GthCBF4* overexpressed plants under cold stress

The transgenic and wild-type *A. thaliana* were grown in plastic bowls for two weeks, the growth pattern between the transgenic and the wild types were not significantly different. The two types, transgenic and the wild type Arabidopsis were placed in an environment of -15°C for 3 hours. After the treatment, the treated plants were moved to a 4°C light incubator to thaw for 4 hours, and finally they were cultivated under normal light conditions at 23°C. After 7 days, photographs were taken to aid the counting of the survival rate of the plants. The germination rate and root length was determined under cold stress; a t-test was used to verify the significance of the difference in root length between the mutant and the wild types. The trypan blue staining method [66] was adopted to reflect the cell damage under cold stress. Moreover, the DAB staining was applied to reflect the accumulation of peroxidase under cold stress. DAB staining was performed using DAB chromogenic kit (Nanjing Jiancheng Bioengineering Institute, Nanjing, China).

Statistical Analysis

The experiments were done in three biological replicates, and the data were statistically analyzed by analysis of variance (ANOVA) procedure [67], using statistical analysis software (SAS) version 9.3. The least significant difference test ($P \leq 0.05$) was used for mean comparison.

Declarations

Ethics approval and consent to participate

Relevant permits/permissions/licences were obtained from relevant institutional, national, and international guidelines and legislation. The permission to carry out this research granted through the Institute of Cotton Research, Chinese academy of Agricultural Sciences (ICR-CAAS), Anyang, China

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Availability of data

All data generated or analysed during this study are included in this article [and its supplementary information files].

Funding

This research was funded by the National Natural Science Foundation of China (32072023), Agricultural Science and Technology Innovation Program of Chinese Academy of Agricultural Sciences

Author Contribution statement

Conceptualization, JL, ROM. and FL.; methodology, JL,ROM, YX,TTW and FL software, JL,ROM, YX,TTW and FL.; validation, JL,ROM, YX,TTW, EO and FL. and Z.Z.; formal analysis, JL,ROM, YX,TTW and FL.; investigation, JL,ROM, YX,TTW and FL.; resources, SAG, KW, ZZ, XC and FL.; data curation, JL,ROM, YX,TTW and FL.; writing—original draft preparation, JL,ROM, YX,TTW and FL.; writing—review and editing, JL,ROM, YX,TTW, JZ, YH, SGA, KW, ZZ, XC and FL.; visualization, JL,ROM, YX,TTW and FL.; supervision, SGA, KW, ZZ, XC and FL.; project administration, SGA, KW, ZZ, XC and FL.; funding acquisition, SGA, KW, ZZ, XC and FL. All authors have read and agreed to the published version of the manuscript.

Acknowledgements

We do appreciate the entire research team of our laboratory for their immense support in the course of this research.

References

1. Hemantaranjan A, Bhanu AN, Singh MN, Yadav DK, Patel PK, Singh R, et al. Heat stress responses and thermotolerance. 2014;1:62–70.
2. Rihan HZ, Al-issawi M, Fuller MP. Advances in physiological and molecular aspects of plant cold tolerance. *J. Plant Interact.* [Internet]. Taylor & Francis; 2017;12:143–57. Available from: <https://doi.org/10.1080/17429145.2017.1308568>
3. Körner C. Plant adaptation to cold climates [version 1 ; referees : 2 approved] Referee Status : 2016;5:1–5.
4. Wang D-Z, Jin Y-N, Ding X-H, Wang W-J, Zhai S-S, Bai L-P, et al. Gene regulation and signal transduction in the ICE–CBF–COR signaling pathway during cold stress in plants. *Biochem.* 2017;82:1103–17.
5. Xiong L, Schumaker KS, Zhu J. Cell Signaling during Cold , Drought , and Salt Stress. *Plant Cell.* 2002;supplement:S165–84.

6. Kumar Verma J, Wardhan V, Singh D, Chakraborty S, Chakraborty N. Genome-wide identification of the alba gene family in plants and stress-responsive expression of the rice alba genes. *Genes* (Basel). 2018;9.
7. Magwanga RO, Kirungu JN, Pu L, Xiu Y, Dong Q, Cai X, et al. Genome wide identification of the trihelix transcription factors and overexpression of *GH_A05G2067 (GT-2)* , a novel gene contributing to increased drought and salt stresses tolerance in cotton. *Physiol. Plant.* [Internet]. 2019;2067. Available from: <http://doi.wiley.com/10.1111/ppl.12920>
8. Diouf L, Pan Z, He S-P, Gong W-F, Jia YH, Magwanga RO, et al. High-Density Linkage Map Construction and Mapping of Salt-Tolerant QTLs at Seedling Stage in Upland Cotton Using Genotyping by Sequencing (GBS). *Int. J. Mol. Sci.* [Internet]. 2017;18:2622. Available from: <http://www.mdpi.com/1422-0067/18/12/2622>
9. Magwanga RO, Lu P, Kirungu JN, Lu H, Wang X, Cai X, et al. Characterization of the late embryogenesis abundant (LEA) proteins family and their role in drought stress tolerance in upland cotton. *BMC Genet.* 2018;19.
10. Lu P, Magwanga RO, Guo X, Kirungu JN, Lu H, Cai X, et al. Genome-Wide Analysis of Multidrug and Toxic Compound Extrusion (MATE) Family in Diploid Cotton, *Gossypium raimondii* and *Gossypium arboreum* and Its Expression Analysis Under Salt, Cadmium and Drought Stress. *G3: Genes|Genomes|Genetics* 2018;8:g3.200232.2018. Available from: <http://g3journal.org/lookup/doi/10.1534 /g3.118.200232>
11. Thomashow MF. Plant cold accliation: Freezing tolerance genes and regulatory mechanisms. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 1999;50:571–99.
12. Bruce TJA, Matthes MC, Napier JA, Pickett JA. Stressful “memories” of plants: Evidence and possible mechanisms. *Plant Sci.* 2007;173:603–8.
13. Colebrook EH, Thomas SG, Phillips AL, Hedden P. The role of gibberellin signalling in plant responses to abiotic stress. *J. Exp. Biol.* 2014;217:67–75. Available from: <http://jeb.biologists.org/cgi/doi/10.1242/jeb.089938>
14. Zhao Q, Xiang X, Liu D, Yang A, Wang Y. Tobacco Transcription Factor NtbHLH123 Confers Tolerance to Cold Stress by Regulating the NtCBF Pathway and Reactive Oxygen Species Homeostasis. *Front. Plant Sci.* [Internet]. 2018;9:1–13. Available from: <http://journal.frontiersin.org/article/10.3389/fpls.2018.00381/full>
15. Liu X, Zhou Y, Xiao J, Bao F. Effects of Chilling on the Structure , Function and Development of Chloroplasts. 2018;9:1–10.
16. Uemura M. Plant plasma membrane proteomics for improving cold tolerance. 2014;
17. Zhao C, Zhang Z, Xie S, Si T, Li Y, Zhu J. Mutational Evidence for the Critical Role of CBF Transcription Factors in Cold Acclimation in Arabidopsis 1. 2016;171:2744–59.
18. Shi Y, Ding Y, Yang S. Molecular Regulation of CBF Signaling in Cold Acclimation. *Trends Plant Sci.* 2018. p. 623–37.

19. Zhang Z, Huang R. Enhanced tolerance to freezing in tobacco and tomato overexpressing transcription factor TERF2/LeERF2 is modulated by ethylene biosynthesis. *Plant Mol. Biol.* 2010;73:241–9.
20. Gong Z, Xiong L, Shi H, Yang S, Xu G, Chao D, et al. Plant abiotic stress response and nutrient use efficiency. 2020;63:635–74.
21. Liu Q, Kasuga M, Sakuma Y, Abe H, Miura S, Yamaguchi-shinozaki K. Two Transcription Factors , DREB1 and DREB2 , with an EREBP / AP2 DNA Binding Domain Separate Two Cellular Signal Transduction Pathways in Drought- and Low- Temperature-Responsive Gene Expression , Respectively , in Arabidopsis. 1998;10:1391–406.
22. Jin Y, Zhai S, Wang W, Ding X, Guo Z, Bai L, et al. Identification of genes from the ICE–CBF–COR pathway under cold stress in Aegilops–Triticum composite group and the evolution analysis with those from Triticeae. *Physiol. Mol. Biol. Plants.* Springer India; 2018;24:211–29. Available from: <https://doi.org/10.1007/s12298-017-0495-y>
23. Jin J-H, Wang M, Zhang H-X, Khan A, Wei A-M, Luo D-X, et al. Genome-wide identification of the AP2/ERF transcription factor family in pepper (*Capsicum annuum* L.). *Genome.* 2018;61.
24. Haake V. Transcription Factor CBF4 Is a Regulator of Drought Adaptation in Arabidopsis. *PLANT Physiol.* 2002;130:639–48.
25. Chinnusamy V, Ohta M, Kanrar S, Lee B ha, Hong X, Agarwal M, et al. ICE1: A regulator of cold-induced transcriptome and freezing tolerance in arabidopsis. *Genes Dev.* 2003;17:1043–54.
26. Ma LF, Zhang JM, Huang GQ, Li Y, Li XB, Zheng Y. Molecular characterization of cotton C-repeat/dehydration-responsive element binding factor genes that are involved in response to cold stress. *Mol. Biol. Rep.* 2014;41:4369–79.
27. Siddiqua M, Nassuth A. Vitis CBF1 and Vitis CBF4 differ in their effect on Arabidopsis abiotic stress tolerance , development and gene expression. 2011;1:1345–59.
28. Yu X, Pijut PM, Byrne S, Asp T, Bai G, Jiang Y. Plant Science Candidate gene association mapping for winter survival and spring regrowth in perennial ryegrass. *Plant Sci.* Elsevier Ireland Ltd; 2015;235:37–45. Available from: <http://dx.doi.org/10.1016/j.plantsci.2015.03.003>
29. Bai GHSML. Signal transduction during cold , salt , and drought stresses in plants. 2011;
30. Bhatnagar-Mathur P, Vadez V, Sharma KK. Transgenic approaches for abiotic stress tolerance in plants: Retrospect and prospects. *Plant Cell Rep.* 2008;27:411–24.
31. Lata C, Yadav A, Pras M. Role of Plant Transcription Factors in Abiotic Stress Tolerance. *Abiotic Stress Response Plants - Physiol. Biochem. Genet. Perspect.* 2011; Available from: <http://www.intechopen.com/books/abiotic-stress-response-in-plants-physiological-biochemical-and-genetic-perspectives/role-of-plant-transcription-factors-in-abiotic-stress-tolerance>
32. Cai X, Magwanga RO, Xu Y, Zhou Z, Wang X, Hou Y, et al. Comparative transcriptome , physiological and biochemical analyses reveal response mechanism mediated by CBF4 and ICE2 in enhancing cold stress tolerance in *Gossypium thurberi*. 2019;1–17.

33. Ambawat S, Sharma P, Yadav NR, Yadav RC. MYB transcription factor genes as regulators for plant responses: An overview. *Physiol. Mol. Biol. Plants*. 2013. p. 307–21.
34. Reddy N, Reddy N, Yang Y. DigitalCommons @ University of Nebraska - Lincoln Extraction and characterization of natural cellulose fibers from common milkweed stems Extraction and characterization of natural cellulose fibers from common milkweed stems. 2009;
35. Chaurasiya PC, Singh S. A Review Report: Low Temperature Stress for Crop Production. 2019;
36. Thakur P, Kumar S. Cold stress effects on reproductive development in grain crops: An overview. 2010;
37. Magwanga RO, Lu P, Kirungu JN, Dong Q, Hu Y, Zhou Z, et al. Cotton Late Embryogenesis Abundant (LEA2) Genes Promote Root Growth and Confers Drought Stress Tolerance in Transgenic *Arabidopsis thaliana*. *G3 Genes, Genomes, Genet*. 2018;
38. Bo H, Longguo JIN, Jinyuan LIU. Molecular cloning and functional characterization of a DREB1 / CBF-like gene (GhDREB1L) from cotton. 2007;50:7–14.
39. Mohseni S, Che H, Djillali Z, Dumont E, Nankeu J, Danyluk J. Wheat CBF gene family : identification of polymorphisms in the CBF coding sequence. 2012;881:865–81.
40. Sunchung P, Ainong S, Mo B. Genome-wide identification and expression analysis of the CBF / DREB1 gene family in lettuce. 2020;1–14.
41. Lv K, Li J, Zhao K, Chen S, Zhang W, Liu G, et al. Plant Science Overexpression of an AP2 / ERF family gene , BpERF13 , in birch enhances cold tolerance through upregulating *CBF* genes and mitigating reactive oxygen species. 2020;292.
42. Marquez-cedillo L, Filichkin T, Amundsen K, Stockinger EJ, Thomashow MF, Chen THH, et al. Structural , functional , and phylogenetic characterization of a large CBF gene family in barley. 2005;533–51.
43. Kidokoro S, Watanabe K, Ohori T, Moriwaki T, Maruyama K, Mizoi J, et al. Soybean DREB1/CBF-type transcription factors function in heat and drought as well as cold stress-responsive gene expression. *Plant J*. 2015;81:505–18.
44. Ofuku KDIJ. The AP2 domain of APETALA 2 defines a large new family of DNA binding proteins in *Arabidopsis*. 1997;94:7076–81.
45. Xia E, Li Y. Genome-Wide Characterization of the C-repeat Binding Factor (*CBF*) Gene Family Involved in the Response to Abiotic Stresses in Tea Plant (*Camellia sinensis*). 2020;11:1–13.
46. Liu S, Lv Z, Liu Y, Li L, Zhang L. Network analysis of ABA-dependent and ABA-independent drought responsive genes in *Arabidopsis thaliana*. 2018;637:624–37.
47. Kim JS, Mizoi J, Yoshida T, Fujita Y, Nakajima J, Ohori T, et al. An ABRE promoter sequence is involved in osmotic stress-responsive expression of the DREB2A gene, which encodes a transcription factor regulating drought-inducible genes in *Arabidopsis*. *Plant Cell Physiol*. 2011;52:2136–46.
48. Wang Q, Guan AEY, Wu AEY, Chen H, Chen AEF, Chu AEC. Overexpression of a rice OsDREB1F gene increases salt , drought , and low temperature tolerance in both *Arabidopsis* and rice. 2008;589–602.

49. Viridi AS, Singh S, Singh P. Abiotic stress responses in plants: roles of calmodulin-regulated proteins. 2015;6:1–19.
50. Gruszka D. DNA damage and repair in plants – from models to crops. 2015;6:1–26.
51. Tondelli A, Francia E, Barabaschi D, Pasquariello M, Pecchioni N. Plant Science Inside the CBF locus in Poaceae. *Plant Sci.* [Internet]. Elsevier Ireland Ltd; 2011;180:39–45. Available from: <http://dx.doi.org/10.1016/j.plantsci.2010.08.012>
52. Miura K, Furumoto T. Cold signaling and cold response in plants. *Int. J. Mol. Sci.* 2013. p. 5312–37.
53. Zarka DG, Vogel JT, Cook D, Thomashow MF. Cold Induction of Arabidopsis CBF Genes Involves Multiple ICE (Inducer of CBF Expression) Promoter Elements and a Cold-Regulatory Circuit That Is Desensitized by Low Temperature 1. 2003;133:910–8.
54. Thalhammer A, Bryant G, Sulpice R, Hinch DK, Planck M, Plant M, et al. Disordered Cold Regulated 15 Proteins Protect Chloroplast Membranes during Freezing through Binding and Folding , But Do Not Stabilize Chloroplast Enzymes in Vivo 1 [W][OPEN]. 2014;166:190–201.
55. Letunic I, Doerks T, Bork P. SMART : recent updates , new developments and status in 2015. 2015;43:257–60.
56. Magwanga RO, Kirungu JN, Lu P, Yang X, Dong Q, Cai X, et al. Genome wide identification of the trihelix transcription factors and overexpression of Gh_A05G2067 (GT-2), a novel gene contributing to increased drought and salt stresses tolerance in cotton. *Physiol. Plant.* 2019;
57. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* 2016;33:1870–4.
58. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, et al. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Mol. Plant.* 2020;
59. Higo K, Ugawa Y, Iwamoto M, Korenaga T. Plant cis-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Res.* 1999. p. 297–300.
60. Horton P, Park KJ, Obayashi T, Fujita N, Harada H, Adams-Collier CJ, et al. WoLF PSORT: Protein localization predictor. *Nucleic Acids Res.* 2007;35.
61. Artico S, Nardeli SM, Brilhante O, Grossi-de-sa MF, Alves-ferreira M. Identification and evaluation of new reference genes in *Gossypium hirsutum* for accurate normalization of real-time quantitative RT-PCR data. 2010;1–12.
62. Hoagland DR, Arnon DI. The water-culture method for growing plants without soil. *Calif. Agric. Exp. Stn. Circ.* 1950.
63. Oa CB, Reyes FC, Sun B, Guo H, Gruis DF, Otegui MS. *Agrobacterium tumefaciens* -Mediated Transformation of Maize Endosperm as a Tool to Study Endosperm. 2020;153:624–31.
64. Dupadahalli K. A modified freeze-thaw method for the efficient transformation of *Agrobacterium tumefaciens*. 2014;3–6.
65. Clough SJ, Bent AF. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. 1999;16:735–43.

66. Berrocal-lobo M. Plant Tissue Trypan Blue Staining During Phytopathogen Infection. 2016;

67. Molugaram K, Rao GS. ANOVA (Analysis of Variance). Stat. Tech. Transp. Eng. 2017.

Figures

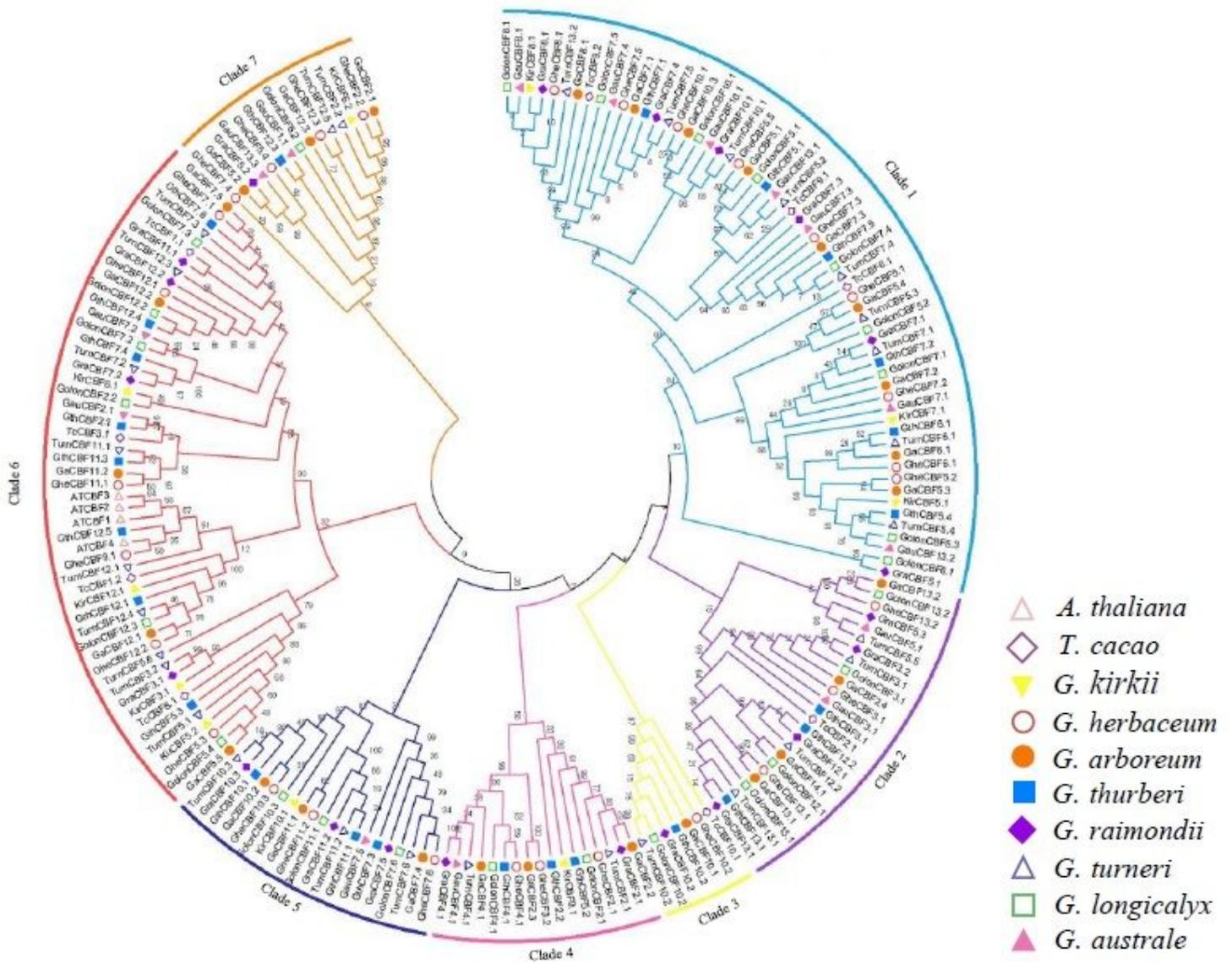


Figure 1

Phylogenetic tree analysis of CBF transcription factors. Protein sequences alignment was done by ClustalX, the phylogenetic tree built by MEG7.0 software using 1000 bootstrap replication via Neighbor-joining approach. Different color was used to differentiate the homologous clusters.

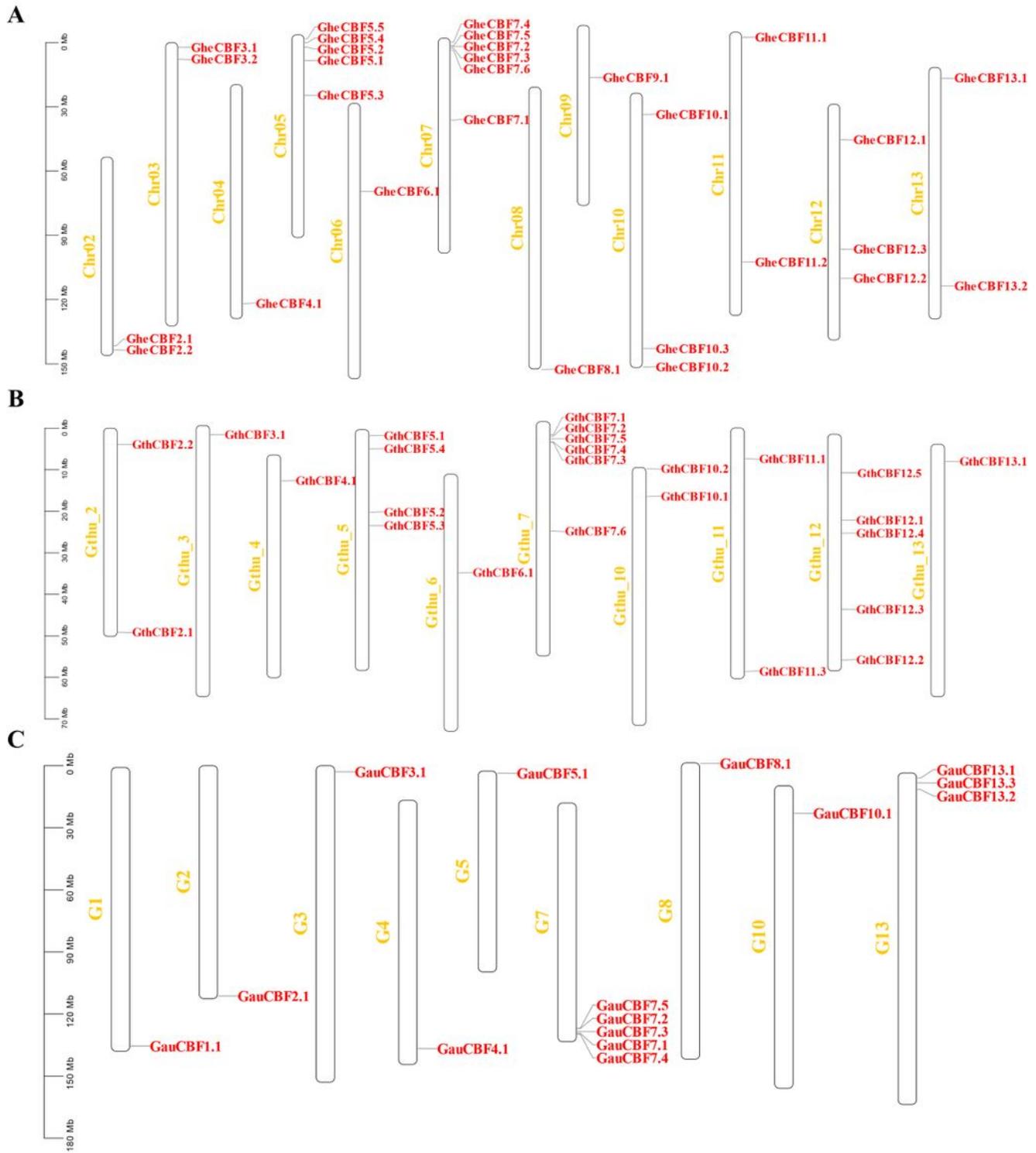


Figure 2

Chromosomal positions of the CBF genes. (A). Chromosomes of *G. herbaceum*, (B). Chromosomes of *G. thurberi* (C). Chromosomes of *G. australe*

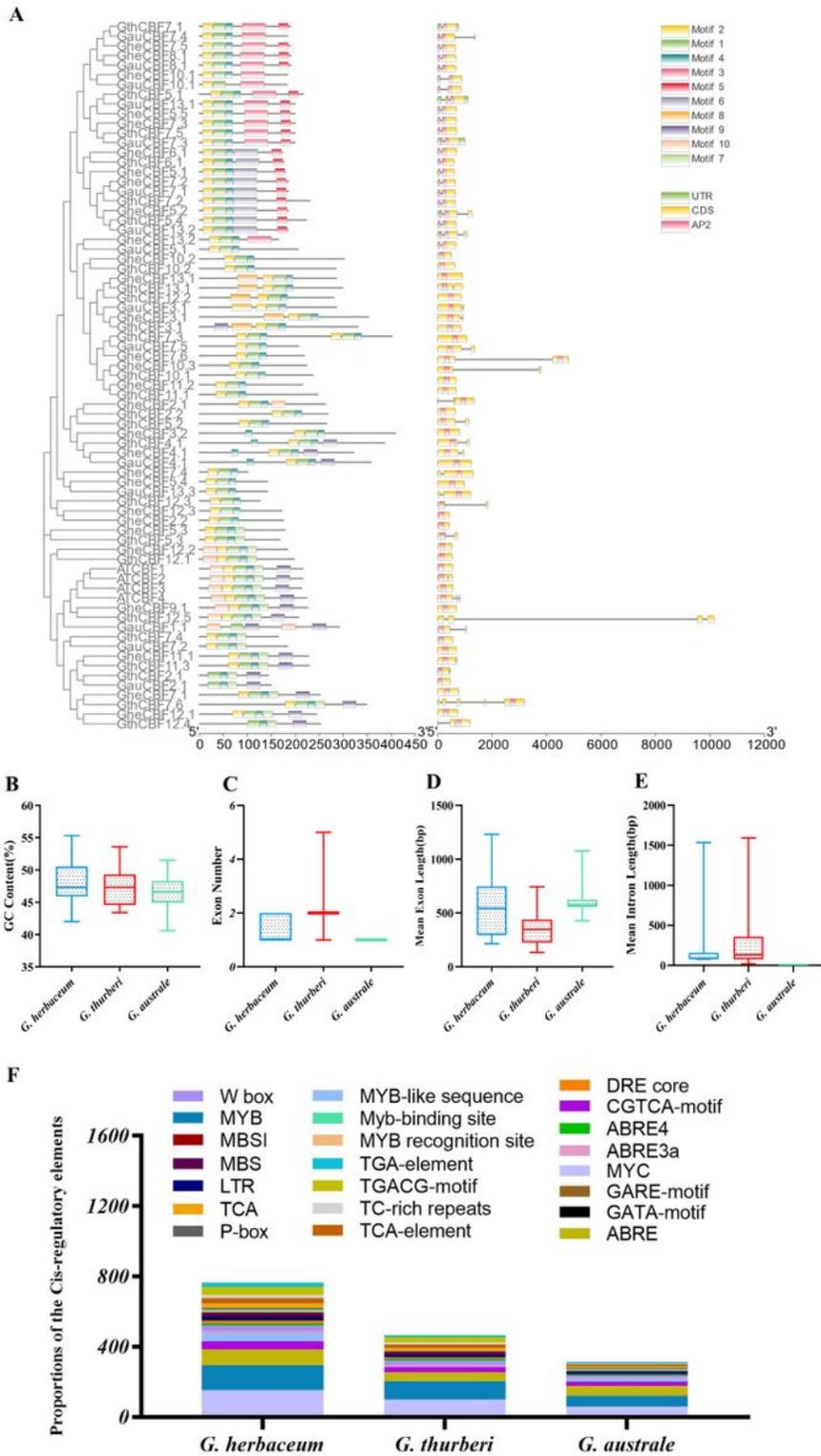


Figure 3

Conserved domain and gene structure analysis of CBF family protein sequences from the evolutionary level in cotton. Different colours denote different motifs. Physiochemical properties and cis-regulatory element analysis of the proteins encoded by the cotton CBF genes. (A-D). Guanine-cytosine (GC) content, exon number, mean exon length and mean intron length. (E). Cis-regulatory elements obtained for the various proteins encoded by the CBF genes in the *G. herbaceum*, *G. thurberi* and *G. australe*.

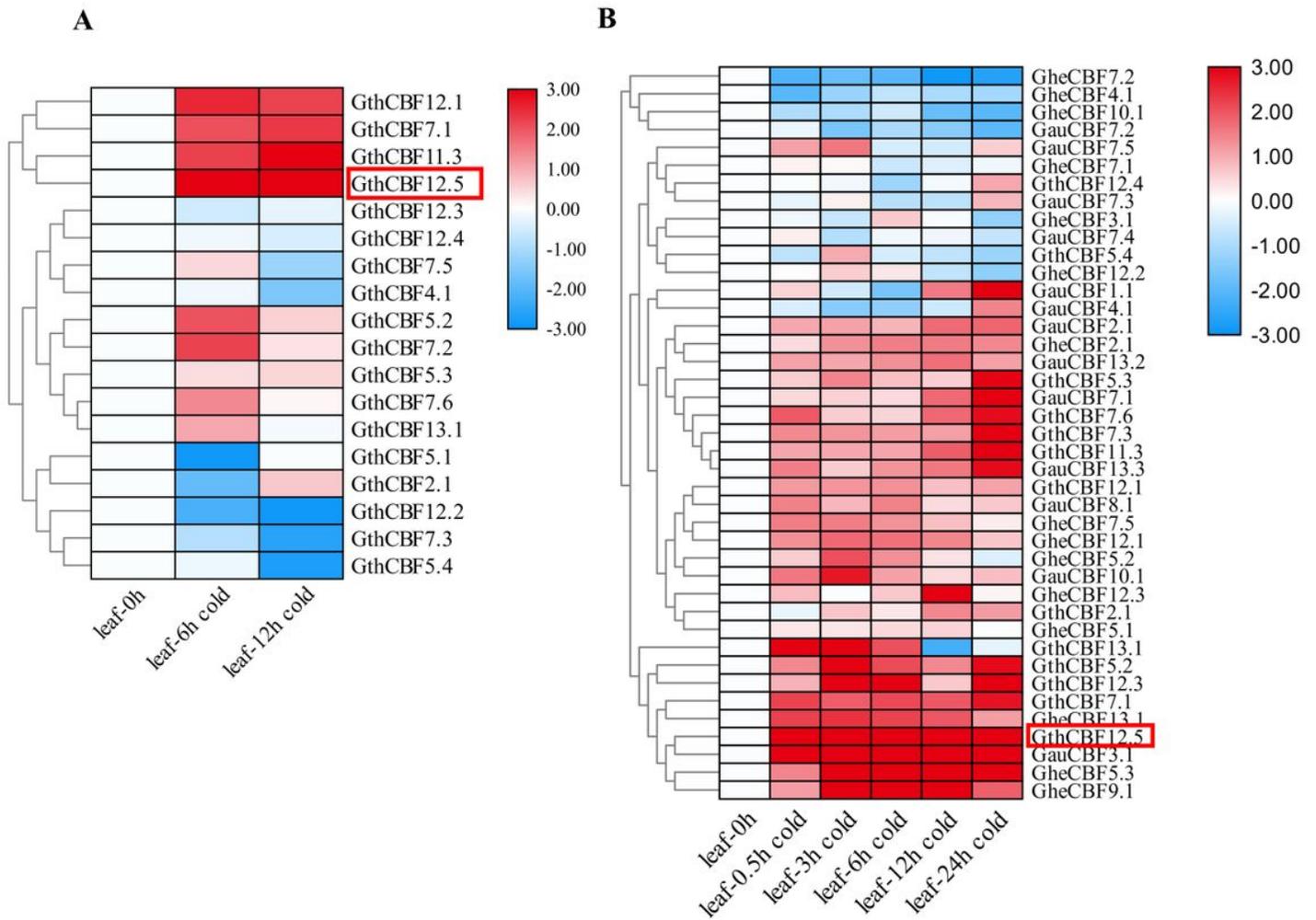


Figure 4

RT-qPCR validations of transcript levels evaluated by RNA-Seq in *G. thurberi* under cold stress. (A) Heat map depicting log₂ (fold change) of transcription factor. (B) Heat map log₂ (fold change) of the CBF genes expression under cold stress condition. The samples were collected at 0h, 0.5h, 3h, 6h, 12h and 24h for leaf tissues. Red: upregulated genes; blue: downregulated; Black: none expressed genes.

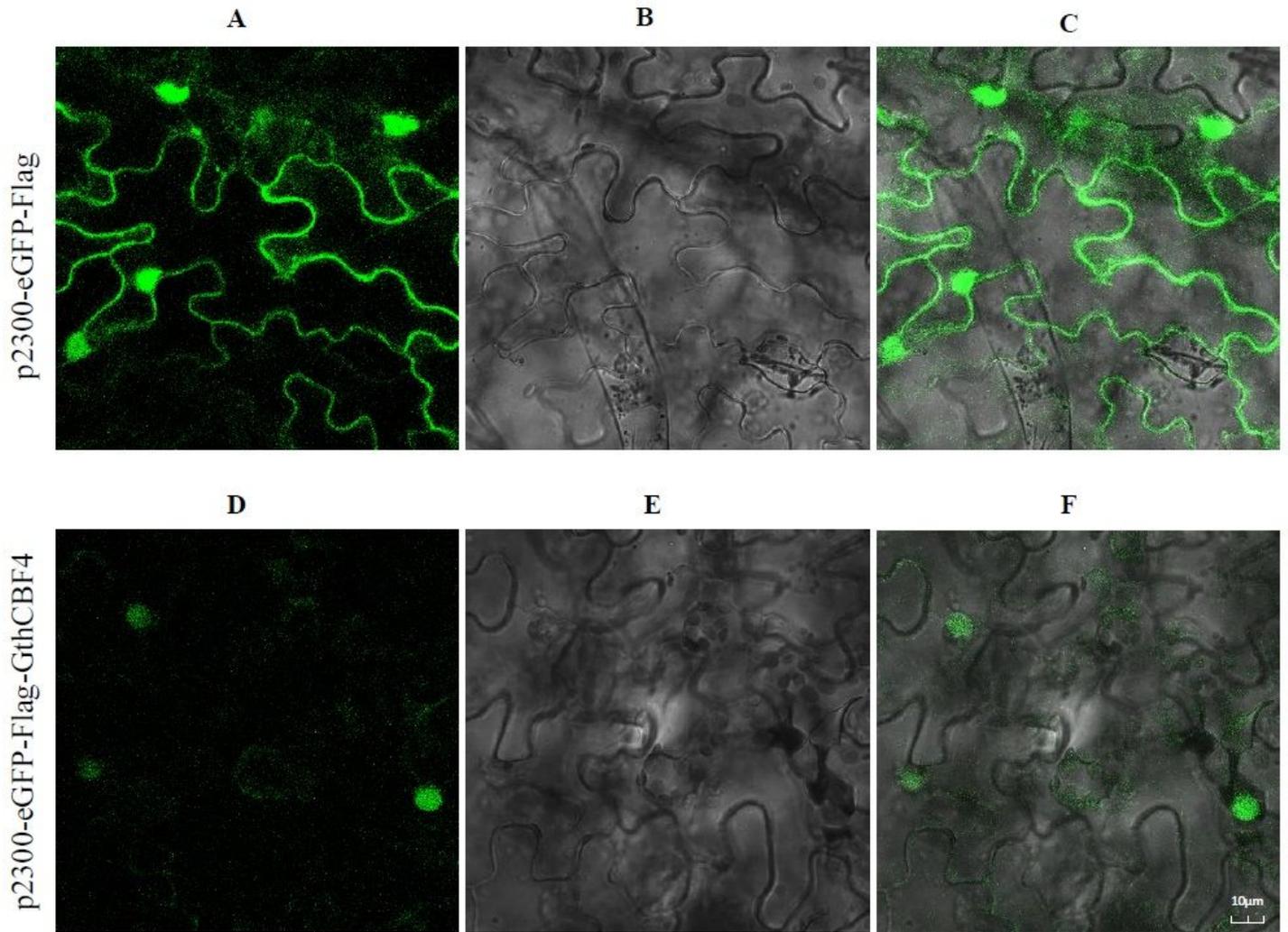


Figure 5

Localization of pCAMBIA2300-eGFP-Flag-GthCBF4 in tobacco leaf. (A–C) Tobacco epidermal cells transformed with pCAMBIA2300-eGFP; (D–F) Onion epidermal cells transformed with pCAMBIA2300-eGFP-Flag-GthCBF4. (A,D) Light field with magnification of X400 to display morphology. (B, E) Dark field images for the detection of green fluorescent protein (GFP) fluorescence. (C, F) Superimposed light and dark field images

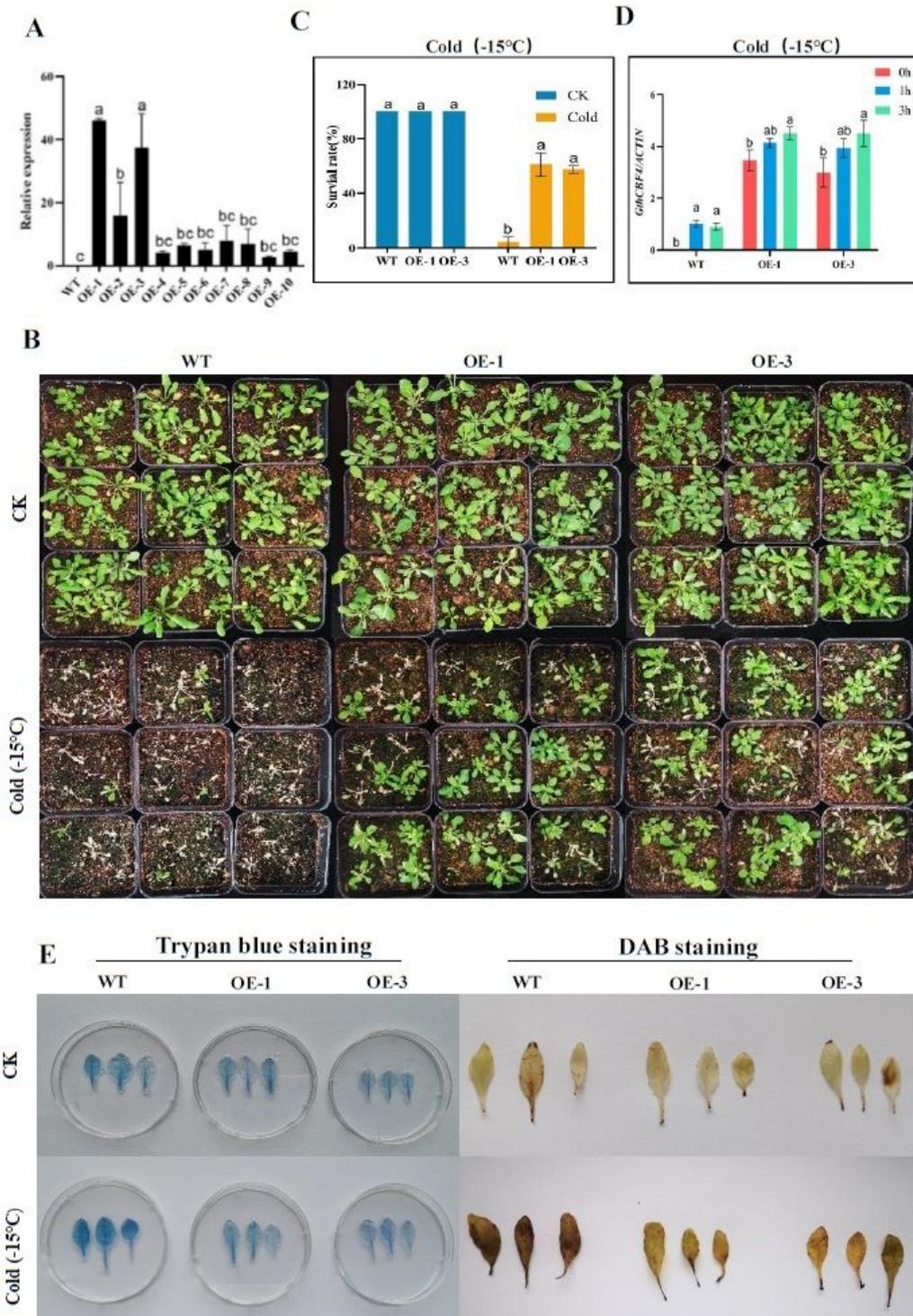
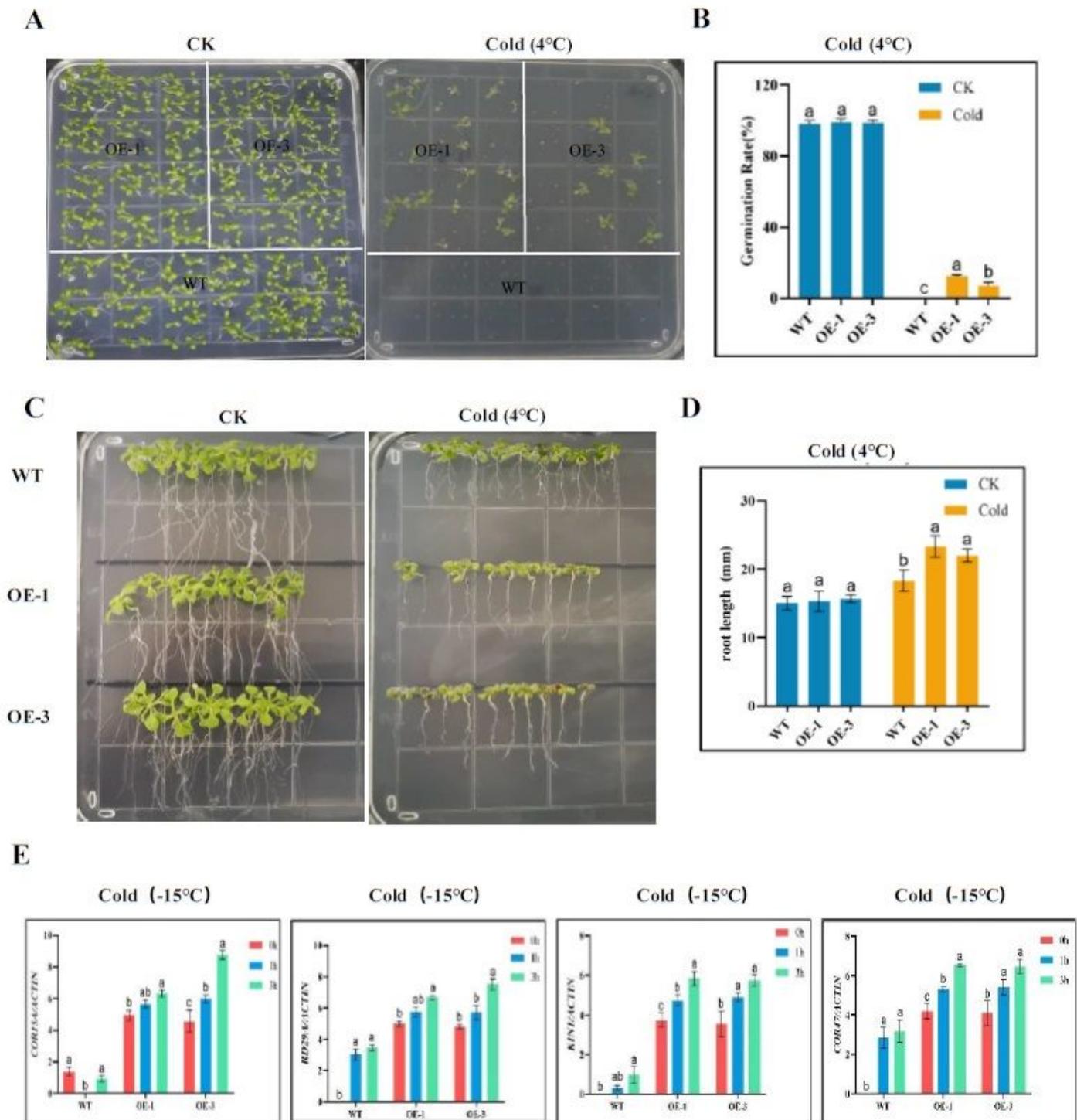


Figure 6

Phenotype identification and cell damage identification of GthCBF4 overexpressed Arabidopsis under low temperature. (A) Quantitative expression level of GthCBF4 gene in transgenic Arabidopsis. (B) -15°C treatment restores the growth status of wild-type and transgenic Arabidopsis thaliana after culture. (C) Freezing survival assay. (D) Log10 (fold change) of the GthCBF4 genes expression in -15°C treatment. (E) Trypan blue staining and DAB staining.



Supplementary Files

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

- [TableS1.docx](#)
- [TableS2.docx](#)
- [TableS3.docx](#)