**Title:** Molecular characterization of three novel ethylene responsive element binding factor genes from cotton

**Running title:** Characterization of three ERF genes from cotton

**Authors:** Long-Guo Jin, Hui Li, Jin-Yuan Liu*

Laboratory of Molecular Biology and MOE Laboratory of Protein Science, Department of Biological Sciences and Biotechnology, Tsinghua University, Beijing 100084, P.R. China

*Author for correspondence. Tel: +86 10 62783845  
Fax: +86 10 62772243  
E-mail: <liujy@mail.tsinghua.edu.cn>

Supported by the National Transgenic Animals & Plants Research Project (2008ZX08009 and 2008ZX08005), the National Natural Science Foundation of China (90608016), the Hi-Tech Research Development Program of China (2007AA02Z141), and China Postdoctoral Science Foundation (2005037052).
Abstract

Ethylene-responsive factors (ERFs) are important regulators of plant gene expression. In this study, three novel ERF genes, \( \text{GhERF2} \), \( \text{GhERF3} \) and \( \text{GhERF6} \), were isolated from cotton (\( \text{Gossypium hirstum} \)) using RACE-PCR. Transient expression analysis using GhERF-GFP fusions showed that these three proteins were targeted to the nucleus. Fusion proteins consisting of GhERF2, GhERF3 or GhERF6 coupled to the GAL4 DNA binding domain strongly activated transcription in yeast. Furthermore, GhERF6 was shown to be able to bind specifically to GCC boxes using a particle bombardment assay in tobacco cells. Semi-quantitative RT-PCR revealed that GhERF2 and GhERF3 are constitutively expressed in all organs, while GhERF6 is only constitutively expressed in vegetative organs. When plants were treated with ethylene, abscisic acid (ABA), salt, cold and drought, the transcripts of GhERF2, GhERF3 and GhERF6 were rapidly induced to high levels. Promoter analysis also indicated that the 5’ upstream regions of the three genes possess elements induced by these physiological and environmental factors. Collectively, our data suggest that GhERF2, GhERF3 and GhERF6 might function as positive \textit{trans}-acting factors in the plant responses to ethylene, ABA and other stresses and provide useful clues for further research into the mechanism of them in regulating cotton multiple stress responses.

\textbf{Keywords:} abiotic stresses; ethylene responsive element binding factors; \textit{Gossypium hirstum}; nuclear localization; transactivation
During their life cycle, plants are often exposed to adverse environmental conditions. Plants have developed complex signaling networks to improve their stress tolerance over the course of evolution. To better understand how plants adapt to various stresses, it is important to explore how different response pathways interact with each other. The cross-talk among these pathways has the potential to regulate and activate multiple stress responses in varying combinations and may help the plant to prioritize the activation of one particular defense pathway over another, thereby providing the optimal defense against a particular abiotic or biotic stress. Transcription factors, through binding to cis-acting elements in the promoters of stress responsive genes or interacting with other transcription factors, subsequently activating or repressing gene expression, play the central role in regulating plant stress response pathways. Studies on some transcription factors have shed light on cross-talk in plant responses to different environmental stimuli (Chen et al. 2004; Fujita et al. 2006; Yamaguchi-Shinozaki et al. 2006). However, the number of transcription factors involved in stress responses is small considering the complex nature of stress response regulation.

The ethylene-responsive factor (ERF) proteins, which make up a large family of transcription factors in plants, are excellent candidates for the study of how plants adapt to various stresses. ERF proteins were reported to modulate the expression of many pathogenesis-related (PR) genes through interaction with a GCC box present in their promoters (Ohme-Takagi et al. 1995; Chakravarthy et al. 2003). In addition, they were also found to bind to non-GCC box cis elements (Chakravarthy et al. 2003). Although only a few members of this family were characterized, such as Pti4, Pti5 and Pti6 from tomato (Gu et al. 2000), NtERF1, NtERF2 and NtERF3 (Ohta et al. 2000) from tobacco and AtERF1 to AtERF5 from Arabidopsis (Fujimoto et al. 2000), most of them have been shown to be crucial in stress and/or hormonal responses.

To further understand the role of ERF proteins in plant stress defense and provide a foundation for applying these genes in enhancement of plant stress resistance, we attempted to clone and characterize this type of gene in an economically important plant species, Cotton (Gossypium hirsutum), which is labile to both abiotic and biotic
stresses. In our previous papers, we reported the characterization and functional analyses of three cotton ERF proteins, GhERF1, GhERF4 and GhERF5 (Qiao et al. 2008; Jin et al. 2008; Jin et al. 2009), which might be functionally important for the acclimation of cotton to stress. In this paper, we report the cloning and expression profiles of three novel cotton ERF genes, GhERF2, GhERF3, and GhERF6, whose expression patterns under various stresses suggest that transcriptional activation cascades involving ERF proteins may be important for plant adaptation to various environmental stresses. Promoter analysis indicated that the 5’ upstream region of these three genes possesses elements induced by some physiological and environmental factors, and these elements may contribute to the defense-related function.

**Results**

**Isolation and sequence analysis of the GhERF2, GhERF3, and GhERF6 cDNA**

To isolate genes from the ERF family of proteins from cotton (*Gossypium hirsutum*), a cloning strategy combining bioinformatics analysis and RACE-PCR was employed. Three cotton ERF-binding protein genes designated GhERF2 (Genbank accession No. AY781117), GhERF3 (Genbank accession No. AY817134) and GhERF6 (Genbank accession No. AY781119) were isolated. GhERF2 has an open reading frame (ORF) of 1173 bp encoding a protein of 390 amino acids with a predicted molecular mass of 43.6 kD; GhERF3 has an ORF of 1191 bp encoding a protein of 396 amino acids with a predicted molecular mass of 43.9 kD; GhERF6 has an ORF of 789 bp encoding a protein of 262 amino acids with a predicted molecular mass of 30 kD (Supplemental Figure S1). Based on homology searches using BLASTX, it was shown that these three GhERF genes displayed high similarities to the genes AtEBP (Y09942) from *Arabidopsis*, LeERF2 (AY192368) and JERF1 (AY044235) from tomato and CaERFLP1 from hot pepper (AY529642). Sequence alignment showed that the three identified proteins and their homologs all had a central 58 amino acid AP2/ERF domain with two functional conserved amino acid residues, alanine at position 14 and aspartic acid at position 19 (Supplemental Figure S1).
S2), which is an identifying characteristic of the ERF binding domain (Sakuma et al., 2002), suggesting that GhERF2, GhERF3 and GhERF6 are new members of the ERF protein subfamily.

To determine the relationship between these three novel GhERF genes with other ERF binding proteins, a phylogenetic analysis based on the amino acid sequence of the AP2/ERF domain was performed, and the results are shown in Figure 1. GhERF2, GhERF3 and GhERF6 were all classified into the B2 subgroup of the ERF subfamily, mainly for possessing a MCGGAI sequence at their N-terminus, which is a conserved motif of the B2 subgroup of ERF subfamily reported previously (Tournier et al. 2003). PSORT server (http://psort.nibb.ac.jp/) analysis indicated that the basic regions KRKRK in GhERF2 and GhERF3 and PAKKRCI in the GhERF6 N-terminal region may function as nuclear localization signals (NLS) and an acidic N-terminal region (corresponding to amino acids 61-73 in GhERF2 and 52–61 in GhERF3) might act as an activation domain for transcription (Supplemental Figure 1). These preliminary results suggest that GhERF2, GhERF3 and GhERF6 might function as transcriptional activators in plants.

Cis-acting element analysis of the promoter regions of the three GhERF genes

To investigate the functions of GhERF2, GhERF3 and GhERF6, the promoter regions of these genes were cloned and a motif search was performed using the PLACE database (http://www.dna.affrc.go.jp/htdocs/PLACE/) to reveal putative cis elements. As shown in Supplemental Figure 1, a number of conserved regulatory motifs present in most eukaryotic promoters was found in the GhERF2 promoter, including ARF (auxin response factor binding site), ARR1 (bacterial response regulator), ASF-1 (involved in transcriptional activation of many genes by auxin and/or salicylic acid), Box-L (involved in the phenylpropanoid biosynthetic pathway), DC3 (DPBP binding core sequence; Dc3 expression is normally embryo-specific and also can be induced by ABA), Dof (signal-responsive and/or tissue-specific gene expression), GARE (GA-responsive element), GT-1 box (involved in pathogen and salt-induced SCaM-4 gene expression), KST1 (target site for Dof1 protein controlling
guard cell-specific gene expression), LTRE (low temperature responsive element), MYB1AT and MYB (recognition site in the promoter of the dehydration responsive gene rd22), SEBF (binding site of the potato silencing element factor SEBF in the PR-10a promoter) and W-box (a binding site of the WRKY transcription factor). This suggests that \textit{GhERF2} may have a complicated role in the regulation of various stress-induced pathways.

The \textit{GhERF3} promoter was also found to possess some elements induced by physiological and environmental factors. These elements may contribute to defense-related functions. For instance, ABA induces ACGT-containing ABRE and DPBF (Hattori et al. 2002). In addition, organ-specific elements, including AACA motifs involved in controlling endosperm-specific expression, the E-box of the napA storage-protein gene, OSE1 and OSE2 organ-specific elements (OSE) and RY/G box and RY repeat elements required for seed-specific expression (Supplemental Figure S1) were also found. These elements may be related to the high level of expression of \textit{GhERF3} in storage tissue.

In the promoter region of the \textit{GhERF6} gene, potential regulatory elements associated with hormone and stress-related responses were found, such as ARR1, CGCG box (a calmodulin-binding/CGCG box DNA-binding protein family involved in multiple signaling pathways), DRE/CRT (the dehydration-responsive element in response to drought, high-salt, and cold stresses), Dc3, GARE, GT-1 box (plays a role in pathogen- and salt-induced SCaM-4 gene expression), MYB and MYC, SEBF, SURE (core of sulfur-responsive element) and W-box elements. The presence of these putative \textit{cis}-acting elements indicates that \textit{GhERF6} might be involved in multiple stress regulation pathways.

\textbf{Expression patterns of three \textit{GhERF} genes}

Analysis of the expression pattern of a new gene can help to reveal its possible biological functions. To investigate the expression patterns of \textit{GhERF2}, \textit{GhERF3} and \textit{GhERF6}, we monitored the level of the corresponding mRNAs in different organs and under stress treatments by semi-quantitative RT-PCR. The results are shown in Figure
We found that *GhERF2* and *GhERF3* had the same expression pattern in different organs, with higher levels of the mRNA transcripts in roots, stems, true leaves and seeds, relatively lower levels in embryos, flowers and fibers, and a very low level in cotyledons. For *GhERF6*, a very different pattern emerged. High levels of *GhERF6* transcript were found in true leaves, stems, roots, slightly lower levels were found in cotyledons, and no *GhERF6* transcripts were detected in young embryos, mature flowers and fibers. This pattern suggests that these three *GhERF* genes are constitutively expressed during vegetative growth under non-stress conditions.

To further characterize whether expression of these three *GhERF* genes is induced by abiotic stresses and exogenous hormones, we investigated the levels of *GhERF* mRNA in 2-week-old seedlings exposed to various treatments (Figure 2). The expression of the three *GhERF* genes could be significantly induced by these treatments, and transcript levels all increased obviously in leaves after 1 h of treatment. Moreover, the induction of the *GhERF2*, *GhERF3* and *GhERF6* transcripts was similar under NaCl, drought and ethylene treatment. After treatment with NaCl, transcripts accumulated rapidly to a maximum after 1 h of exposure and remained unchanged for up to 24 h. In contrast, drought induced *GhERF* expression at a relatively lower levels, with transcripts reaching maximum levels 1 h after treatment and maintaining this level for up to 6 h, then declining gradually. The expression levels of these three genes clearly increased 3 h after spraying with ethephon (1 mM) and reached their peak levels at 12 h, after which the bands became smeared. After treatment with cold, the transcripts of *GhERF2* and *GhERF6* were rapidly induced, reached a maximum at 3 h and remained unchanged up to 24 h. *GhERF3* transcripts also reached the maximum at 3 h, but declined gradually after 6 h. In addition, the expression levels of *GhERF3* and *GhERF6* were relatively lower than that of *GhERF2* at the same time point. When treated with ABA, the transcripts of *GhERF2* and *GhERF6* were significantly induced after 1 h and maintained the same levels up to 24 h, but the transcripts of *GhERF3* reached a maximum at 1 h after treatment, maintained this level until 6 h, and then declined gradually.

The hormones and abiotic stresses tested in this study all obviously induced the
expression of the GhERF2, GhERF3 and GhERF6 genes. This suggests that these three GhERF genes may be involved in the regulation of the response to these stresses.

Nuclear localization assay

Analysis of the amino acid sequences of GhERF2, GhERF3 and GhERF6 revealed that they all contain one nuclear localization sequence (NLS) in their N-terminal region: KRKRK in GhERF2 and GhERF3 and PAKKRCI in GhERF6 (Supplemental Figure 1). To validate the nuclear localization of these genes, the coding regions of GhERF2, GhERF3 and GhERF6 were fused in-frame with the green fluorescence protein (GFP) gene, and the resulting constructs were then bombarded into the epidermal cells of onion. As shown in Figure 3, the fluorescence of GhERF2-GFP, GhERF3-GFP and GhERF6-GFP were all localized exclusively to the nucleus, whereas the fluorescence of GFP alone was observed in the whole cell. This clearly demonstrated that they are nuclear-localized proteins.

Transactivation assay

To investigate the transcriptional activity of GhERF2, GhERF3 and GhERF6, we performed a transient expression assay using a GAL4-responsive reporter system. As shown in Figure 4, the three effector plasmids contain translational fusions between the GAL4 binding domain coding region and GhERF2, GhERF3 or GhERF6. If these GhERF proteins have transactivation activity, after the corresponding effector plasmids are transformed into the yeast strain AH109 harboring two reporter genes, lacZ and His3, the expression of the reporter genes will be activated, resulting in successful growth of transformed yeast cells on SD/-His plates supplemented with 5 mM 3-AT and lacZ. Additionally, the yeast colonies will be blue. Figure 4 shows that the transformed yeast cells harboring pBD-GhERF2, pBD-GhERF3, pBD-GhERF2 and pCL1 (positive control) could grow on the medium and formed blue colonies. In contrast, the negative control yeast cells containing only the pGBKT7 vector did not grow on the same plate. These results indicate that GhERF2, GhERF3 and GhERF6
can activate the expression of the two reporter genes and are active transcriptional activators in yeast.

The ERF proteins have been reported to modulate the expression of many pathogenesis-related (PR) genes through interaction with the GCC box present in their promoters. In order to determine whether these novel GhERF genes can activate GCC-dependent transcription, the GhERF6 was selected to perform a transient expression assay in tobacco cells. The GCC elements were repeated four times in tandem and were fused to a minimal TATA box with the \( LUC \) gene. This construct was then used as reporter plasmid (Figure 5A). The effector construct consisted of the full-length cDNA of \( \text{GhERF6} \) fused to the CaMV 35S promoter. This effector plasmid, along with the 4XGCC-LUC reporter construct, was transferred to mature tobacco leaf cells by particle bombardment. The \( \text{GhERF6} \) gene produced an approximately 7.6-fold increase in LUC activity over the control (Figure 5B). This shows that GhERF6 is able to bind specifically to GCC upstream promoter elements and activate \( LUC \) gene expression.

**Discussion**

The ERF family is a large family of plant transcription factors and is part of the AP2/ERF superfamily, which also contains the AP2 and RAV families (Riechmann et al. 1998). ERF proteins were identified and implicated in a variety of cellular processes such as hormonal signal transduction (Ohme-Takagi et al. 1995), response to biotic (Gu et al. 2000; Yamamoto et al. 1999) and abiotic stresses (Liu et al. 1998; Dubouzet et al. 2003), regulation of metabolism (Van der Fits et al. 2000; Zhang et al. 2005) and in developmental processes (Banno et al. 2001; Chuck et al. 2002). However, our understanding of the plant ERF proteins is still limited as most members of the ERF family have yet to be studied, even in the model plants \( \text{Arabidopsis} \) and rice, despite the likelihood that these genes play important roles in the regulation of many physiological processes (Nakano et al. 2006; Hu et al. 2008). A great deal of experimental work will be required to determine the specific biological function of each of these genes. In this study, three novel members of this
subfamily, *GhERF2*, *GhERF3* and *GhERF6* were isolated from cotton. Based on the conserved 14th alanine and 19th aspartic acid residues in the AP2/ERF domain and the N-terminal MCGGAII/L signature sequence, they were classified as members of the B2 group of the ERF subfamily, which has been shown to play an important role in ethylene and pathogen response and has multiple functions in the regulation of GCC-mediated gene expression in plants (Tournier et al. 2003; Lee et al. 2004). Fusion proteins containing *GhERF2*, *GhERF3* or *GhERF6* coupled to the GAL4 DNA binding domain strongly activated transcription in yeast. These features, in combination with their nuclear localization, demonstrate that *GhERF2*, *GhERF3* and *GhERF6* are active transcription factors.

To provide some insight into the biological roles of these three novel *GhERF* genes, we first analyzed the conserved cis-acting elements in the promoter regions. Preliminary analysis indicated that potential regulatory elements associated with hormone and abiotic stress-related responses are present in the promoter regions of these *GhERF* genes, such as ACGT-containing ABRE and DPBF induced by ABA; A/GCCGAC-containing DRE regulating gene expression in response to drought, high-salt, cold and ABA; and LTRE (low-temperature responsive element). These elements might contribute to the corresponding defense-related functions. However, promoter analysis only provides some clues regarding the mechanism by which these genes respond to stress signals. The functional significance of these putative regulatory elements for the expression of *GhERF* genes in cotton needs to be elucidated.

Analysis of the expression patterns of *GhERF2*, *GhERF3* and *GhERF6* in response to hormones and other abiotic stresses was performed. The expression of the three *GhERF* genes was found to be significantly induced by cold, NaCl, drought, ethylene and ABA treatment, and the transcript levels all increased obviously in leaves after 1 h of treatment. Previous work showed that plant stress responses were regulated by multiple signaling pathways and that there was significant overlap of the genes expressed in response to different stresses (Singh et al. 2002). In addition, plant hormones including ethylene and ABA play a role in these processes (Finkelstein et al. 2003).
2002; Wang et al. 2002). For example, ABA mediates the response of plants to drought, cold, and salt stress (Zhu 2002). The observation that both ethylene and ABA can induce the expression of these three GhERF genes suggests that they might have a role in the cross-regulation of the cotton plant response to these two hormones. Previous studies showed that many other ERF genes could also be induced by exogenous hormones and abiotic stresses. For example, the mRNA transcripts of AtERF1, AtERF2 and AtERF5 increased 2- to 3-fold after 12 h of ethylene treatment (Fujimoto et al. 2000). The expression of the CaERFLP1 gene was induced by treatment with ethylene and high salinity, and was only slightly affected by drought and cold treatment (Finkelstein et al. 2002). The expression of tomato JERF1 and JERF3 was also induced by ethylene, ABA, salt and/or cold treatments (Zhang et al. 2004). These results strongly suggest that ERF genes in different plant species all might play important roles in the regulation of plant stress defense through different signaling pathways.

In this study, it was shown that GhERF2, GhERF3 and GhERF6 may act as important common components of multiple signaling pathways responsive to biotic and abiotic stresses. To further understand their function, the development of transgenic plants over-expressing of these genes is underway.

Materials and Methods

Plant growth and stress treatments

Cotton (Gossypium hirsutum cv. Zhongmian 12) plants were grown in potting soil in a culture room at 25 °C with a 16 h light and 8 h dark cycle. Abiotic stress treatments were applied to 2-week-old seedlings for 24 h or for the specific time where indicated. Low temperature treatments were applied by transferring plants to a growth chamber set to 4 °C for different periods of time under the light and photoperiodic conditions described above. Dehydration was performed by removing plants from the pots and placing them on a dry filter paper. The relative water content of dry-treated seedlings was reduced to 50%. Salinity and ABA treatments were applied by submerging the roots of the plants in an aqueous solution of 400 mM NaCl
or 100 μM ABA. Ethephon, which emits ethylene when dissolved in water, was used as a substitute for ethylene. During the treatments, the Petri-dish was sealed with parafilm and then the whole plants were harvested 1, 3, 6, 12 and 24 h after treatment. All tissues harvested for nucleic acid extraction were weighed, immediately frozen in liquid nitrogen, and stored frozen at -70 °C until use.

Isolation of the full length cDNA of *GhERF2*, *GhERF3*, and *GhERF6* by RACE-PCR

Total RNA was isolated from seedlings and various organs using the RNeasy Plant Mini Kit (Qiagen, CA, USA) following the manufacturer’s instructions. The cotton EST database (http://www.tigr.org/) was searched using the blastn program with the nucleotide sequence of AtEBP (Y09942) (Büttner et al. 1997), and three cotton EST clones were found with a predicted protein sequence highly homologous to that of the AP2 domain of the ERF transcription factor. These three ESTs were then used as templates to design gene-specific primers. The 5’ and 3’ RACE kits (Takara, Dalian, China) were employed to isolate full-length cDNAs following the manufacturer’s instructions. For 3’ race, the oligonucleotides (3sites Adaptor Primer) included in the 3′-RACE kit were used as the antisense primers, and the two gene-specific primers 3′-GSP3 (5′-AGGATGAGGATTCTGATGTCGA-3′), 3′-NGSP3 (5′-ATGATGTTTTGGCTGATGTTAAGTC-3′) and 3′GSP6 (5′-GAGTTTCCAGACTGACTCACCAC-3′), 3′NGSP6 (5′-GAGAGTGAGTTGGAGTTGAAAGAA-3′) were used as the nested sense primers to isolate the 3′ ends of *GhERF3* and *GhERF6*, respectively, and P1 (5′-CGACTTGGGCGATGATTTTCGAGAC-3′) was used for *GhERF2*. For 5′ race, 5′-phosphorylated RT-primer (5′-PO₄-CGATACTGGTTTCTTC-3′) was used to synthesize single-stranded cDNA, P1 (5′-CGACTTGGGCGATGATTTTCGAGAC-3′) and P2 (5′-ATCTCTTACCCGACCCCTTCTTC-3′) were used for inverse PCR to isolate the 5′ end of *GhERF2*, and F1 (5′-ACCAACCCATTATTTCTACACTACC-3′) and F2 (5′-CATAGACAAATTCAGACACAGTTA-3′) were used to obtain the full-length cDNA of *GhERF2* using a cotton cDNA library as the template; 5′-GSP3
(5′-CTTCGAGTGCTTTCTACCCGAC-3′), 3′-GSP3 and 3′-NGSP3, 5′-NGSP3
(5′-AATCGGAGATAATCGCACCTC-3′) for end-to-end PCR to isolate GhERF3;
5′GSP6 (5′-GAGTAAGAGCTGGTGAAGCTGCT-3′), 3′GSP6 were used for first
round PCR and 5′NGSP1 (5′-CGAGTTGCTGAGTGCTTTCAC-3′), 3′NGSP1
were used for nested PCR to isolate GhERF6.

Sequence alignment and phylogenetic analysis

DNA sequence data were assembled and analyzed using the DNAMAN analysis
program (Lynnon Biosoft, USA). Database searches were performed with the
National Center for Biotechnology Information (Bethesda, MD USA) BLAST search
program. Alignment of the structurally related AP2/ERF proteins was performed using
Clustal X software (Thompson et al. 1994) and further adjusted using the GenDoc
program (Nicholas et al. 1997). A phylogenetic tree was drawn with the TreeView
program (Page 1996).

Isolation of the Promoter region through genome walking

Total genomic DNA was extracted from leaf tissue as described by Paterson
(Paterson et al. 1993) and digested with the proper restriction endonucleases. Cloning
of the promoter region was performed with an improved PCR-based genomic walking
method described previously (Wu et al. 2006). Major PCR band(s) were isolated from
the gel using the Qiagen Gel extraction III kit (Qiagen, Germany), and the isolated
fragments were then cloned into a pMD18-T vector (TakaRa, Dalian, China).
Recombinant plasmid DNA used for sequencing was prepared using the QIAprep
Spin Mini Prep Kit (Qiagen, Germany), and the inserts were sequenced using a
BigDyeTM Terminator Cycle Sequencing Ready Reaction Kit (Perkin Elmer, USA)
on an ABI PRISMTM 377 DNA Sequencer. The promoter sequences were analyzed
with the PLACE database (http://www.dna.affrc.go.jp/htdocs/PLACE/) (Higo et al.
1999). The PCR-amplified products were cloned into the pMD18-T vector (TakaRa,
Dalian, China) and sequenced.
Semi-quantitative RT-PCR analysis

Total RNA was isolated from different parts of the mature plants or seedlings using the RNeasy plant mini kit (Qiagen, CA, USA) following the manufacturer’s instructions. RNA concentration was estimated based on absorbance at 260 nm. Semi-quantitative RT-PCR analysis was carried out as described previously (Huang et al. 2006a). The gene-specific primers were as follows:

RS2(5′-CCATCGGTCAATCAGTTCGCATAC-3′) and RA2(5′-CAGTGCAGGCAAGGCTTAAGC-3′) for GhERF2; RS3(5′-AGCATGCTGTAAAATGAATTCTC-3′) and RA3(5′-ACAATAGTCAAGCAGGCTTAAGC-3′) for GhERF3; RS6(5′-AGCAGCTTCACCAGCTCTTACTC-3′) and RA6 (5′-GTTAGCGCCGGGCTTAAGC-3′) for GhERF6; and SSU1 (5′-AACCTTAAAGGAATTGACGGAAG-3′) and SSU2 (5′-GCATCACAGACCTGTGATGG-3′) for cotton small-subunit (SSU) rRNA. The cotton small-subunit rRNA was used as the control to normalize the amount of template in the PCR reactions. In order to ensure the gene-specificity of these primers, we sequenced the PCR fragments obtained with these primers and found they were indeed derived from the three GhERF transcripts or cotton small-subunit (SSU) rRNA, respectively. The PCR products were then separated on a 1.2% (w/v) agarose gel. All RT-PCR expression assays were performed and analyzed at least three times in independent experiments.

Subcellular localization analysis of transiently expressed fusion proteins

The coding regions of the GhERF2, GhERF3 and GhERF6 genes were amplified by PCR and inserted into the Nco I and Bam HI sites of the pCK-GFP vector to generate GhERF-GFP in-frame fusion proteins. The fusion constructs or the control GFP vector were introduced into onion epidermal cells by particle bombardment using a Bio-Rad Biolistic PDS 1000/He system. Plasmid DNA (1 μg) was used to coat 0.5 mg of 1 μm gold particles as described by Huang (Huang et al. 2006b). The onion
epidermal cells had previously been incubated on MS agar plates in the light at 22 ºC for 24 h. The initial pressure of bombardment was 1100 psi, and the traveling distance of the particles to the plant tissues was 6 cm. Bombarded tissues were placed on the same agar plates and incubated at 22 ºC for 24 h in the dark, followed by monitoring the localization of GFP with a confocal microscope (Olympus, FluoView™ FV300).

**Transactivation assay based on the yeast GAL4 system**

Transactivation by the three *GhERF* genes was examined in a strain of *Saccharomyces cerevisiae*, AH109, in which the *lacZ* and *His3* reporter genes are driven by the *GAL1* and *MEL1* promoters, respectively, and in which the transformation markers *trp1-901* and *leu2-3* are present. Effector plasmids were constructed by ligating the entire coding region of each *GhERF* gene into the *BamHI/PstI* sites of the pGBKT7 vector (Clontech, USA) to produce the fusion protein with GAL4 DNA binding domain (BD). The resulting constructs (pGBKT7-GhERF) and the empty control vector (pGBKT7) were transformed into AH109 yeast cells according to the manufacturer’s instructions (Clontech, USA). The vector pCL1 encodes the full-length, wild-type GAL4 protein and was used as a positive control. The transformants were selected by growing on Trp-synthetic dropout medium at 30 ºC for 3 days. The β-galactosidase activity was assayed on synthetic dextrose (SD) plates supplemented with 5 mM 3-AT (3-aminotriazole) to suppress leaky histidine expression and without His according to the method of Duttweiller (Duttweiller et al. 1996).

**Transactivation of the 4×GCC-LUC reporter genes in tobacco cells**

Transient transactivation assays were performed on tobacco cells using the particle bombardment method. The reporter plasmid 4×GCC-LUC (Ohta et al., 2000) was kindly provided by Dr. Masaru Ohme-Takagi (National Institute of Advanced Industrial Science and Technology, Tsukuba, Japan). The effector plasmid 35S-GhERF6 was constructed by replacing the GUS gene in pBE2113-GUS (Mitsuhara et al. 1996) with the coding region of the *GhERF6* gene. A blank plasmid
of pBE2113-GUS without the GUS gene was used as the negative control. The plasmid pRTLNLUC (CaMV35S-Renilla LUC-NOS) provided by Dr. Gregory B. Martin (Cornell University, USA) was used as a reference plasmid. Transient assays were performed using the particle bombardment method (Sanford et al. 1993). After bombardment, the samples were incubated on filter paper moistened with 50 mM phosphate buffer (pH 7.0) in a dark chamber at 22 ºC for 20 h. The samples were then frozen in liquid nitrogen for luciferase assays. LUC assays were performed with the dual-luciferase reporter assay system (Promega, USA) using a luminescence reader (TD-20/20, Turner, USA). To normalize values after each assay, the ratio of LUC activities (Firefly LUC/Renilla LUC) was calculated. The data were collected from triplicate samples and three independent experiments.

Acknowledgments

We are grateful to Dr. Gregory B. Martin (Cornell University, USA) for providing us with the p35S-LUC and pRTL-NLUC plasmids and Dr. Masaru Ohme-Takagi for providing us with the reporter plasmid 4xGCC-LUC. This work was supported by grants from the National Natural Science Foundation of China (90608016), the Hi-Tech Research and Development Program of China (Grant No. 2007AA02Z100), and the National Transgenic Animals & Plants Research Project (2008ZX08009-003 and 2008ZX08005-003).

References


Thompson JD, Higgins DG, Gibson TJ (1994). CLUSTAL W: improving the


**Figure Legends:**

Figure 1. Phylogenetic analysis of GhERF6 and other plant EREBP transcription factors. AP2/ERF proteins were aligned using the CLUSTALX program, and the phylogenetic tree was drawn using the TreeView program. The GenBank accession numbers of all proteins used here are: AtEBP(Y09942), AtERF1(AB008103), AtERF2(AB008104), AtERF3(AB008105), AtERF4(AB008106), AtERF5(AB008107), DREB1A(AB007787), DREB1B(AB007788), DREB1C(AB007789), DREB2A(AB007790), DREB2B(AB007791), RAP2.1(NM_103607), RAP2.4(NM_106457), RAP2.6(NM_103468), RAP2.9(NM_179009), RAP2.10(NM_119854), RAP2.11(NM_121984), TINY(NM_122482), ABI4(NM_129580), CBF4(NM_124578), ERF1(EU395634), GhERF1(AY181251), GhERF2(AY781117), GhERF3(AY817134), GhERF4(AY781120), GhDREB1L(DQ409060), GhDBP1(AY174160), GhDBP2(AY619718), GhDBP3(DQ224382), ZmDBF1(AF493800), ZmDBF2(AF493799), and ZmABI4(AY125490).

Figure 2. Expression profiles of **GhERF2**, **GhERF3** and **GhERF6**.

(A) Organ-specific expression of **GhERF2**, **GhERF3** and **GhERF6** in cotton. mRNA was isolated from seeds (SD), roots (RT), stems (ST), cotyledons (CL), true leaves (TL), flowers (FL), embryo (EB) and fiber (FB).

(B) Expression profiles of **GhERF2**, **GhERF3** and **GhERF6** after various treatments including cold, drought, NaCl, ABA and ethylene. The cotton small-subunit rRNA (SSU) was used as the control to normalize the amount of template in the PCR.

Figure 3. Nuclear localization of GhERF2-GFP, GhERF3-GFP and GhERF6-GFP in onion cells. Onion epidermal cells were transformed with 35S-GFP (A-C), GhERF2-GFP (D-F), GhERF3-GFP (G-I), and GhERF6-GFP (J-L). After incubation for 24 h, the transformed cells were observed under a confocal microscope. The photographs were taken in the dark field for green fluorescence (A, D, G, J), in the
bright light for the morphology of the cells (B, E, H, K) and in combination (C, F, I, L).

Figure 4. Transactivation activity of GhERF2, GhERF3 and GhERF6 in yeast cells.

(A) Schematic diagram of the effector and reporter constructs used in the yeast assays. The effectors contained the GAL4 DNA-binding domain coding region (GAL4BD) fused to GhERF2, GhERF3 or GhERF6. The reporter genes were the *LacZ* gene or the *HIS3* gene.

(B) The transformants were selected by growth on Trp− synthetic medium at 30°C for 3 d. The existence of transcriptional activation activity was confirmed by β-galactosidase assay on SD/-His plates supplemented with 5 mM 3-AT.

Figure 5. Transient expression assays in tobacco cells.

(A) Schematic overview of the effector, reporter and reference plasmid constructs used in the particle bombardment experiments. The effector plasmids contain the CaMV 35S promoter fused to full-length GhERF6 cDNA. The reporter constructs possess the GCC element tandemly repeated four times, which were fused to the reporter *luciferase* gene. The reference plasmid is composed of the CaMV 35S promoter and the *luciferase* gene. Nos depicts the polyadenosyl signal of the *nopaline synthase* genes.

(B) Transactivation of the 4XGCC-LUC fusion genes by the GhERF6 protein in tobacco cells. Reporter genes containing GCC upstream *cis*-acting elements were transfected by bombardment with the effector plasmid or vector control in wild-type tobacco leaf cells. To normalize transfection efficiency, the plasmid carrying the CaMV 35S promoter-*luciferase* gene was co-bombarded in each experiment. LUC activity was expressed in arbitrary units relative to luciferase activity. This assay was repeated three times with similar results. The values are means ± SD (n = 3).
Figure 1

ERF subfamily

B-1
AERF3
AERF4
GhERF1
GhERF2
GhERF3

B-2
AERF3
GhERF4
GhERF1

B-3
AERF4
GhERF2
GhERF1

B-4
RAP2.6

B-5
AL035394
AC016997

A-1
UNDREB2L
CBF4
DREB1C
DREB1A
DREB1B

A-2
DREB2B
DREB2A

ZmAB4
AB4

A-3

A-4
GhDBP2
TINY
ZmDBF2

A-5
RAP2.10
GhDBP1

0.1

DREB subfamily
Figure 2
Figure 3
(A) Schematic diagram of plasmids and reporter genes

Effector plasmids

- pGBK7
- PCL1
- pGBK7-GhERF2
- pGBK7-GhERF3
- pGBK7-GhERF6

Wild type Gal4

GAL4BD

GhERF2

GAL4BD

GhERF3

GAL4BD

GhERF6

Yeast Strain AH109 reporter genes

- GAL1 UAS
- GAL1 TATA
- HIS3
- MEL1 UAS
- MEL1 TATA
- LacZ

(B)
Figure 5

(A)

Reporter Plasmid

Effector Plasmid

Control Plasmid

(B)

Control

GhERF6

Relative Luciferase Activity