ISSN 0006-2979, Biochemistry (Moscow), 2006, Vol. 71, No. 3, pp. 285-293. © Pleiades Publishing, Inc., 2006.
Published in Russian in Biokhimiya, 2006, Vol. 71, No. 3, pp. 361-371.
Originally published in Biochemistry (Moscow) On-Line Papers in Press, as Manuscript BM05-201, December 18, 2005.

Cloning and Characterization of Two EREBP Transcription Factors from Cotton (*Gossypium hirsutum* L.)

Hongying Duan[1], Fuguang Li[2], Xudong Wu[3], Dongmei Ma[3], Mao Wang[1], and Yuxia Hou[3]*

[1]*College of Biological Sciences, China Agriculture University, Beijing 100094, P. R. China*
[2]*Cotton Research Institute, Chinese Academy of Agricultural Sciences, Henan Anyang 455004, P. R. China*
[3]*College of Science, China Agriculture University, Beijing 100094, P. R. China;*
fax: 86-10-627-33824; E-mail: houyuxia@tsinghua.org.cn

Received September 5, 2005
Revision received October 10, 2005

Abstract—In this research, two EREBP (ethylene response element binding protein) genes were isolated by the yeast one-hybrid system, named *GhEREB2* and *GhEREB3*, and both have one intron in their coding regions. The deduced amino acid sequences of GhEREB2 and GhEREB3 have some typical features of transcription factors, one potential basic nuclear-localization signal, one possible acidic activation domain, and one conserved DNA binding domain, and they show high similarity, especially in the DNA-binding domain. *GhEREB2* was expressed in roots, stems, and leaves (low levels), but *GhEREB3* was only expressed in leaves. Furthermore, the expression of *GhEREB2* and *GhEREB3* was induced by ethylene and jasmonic acid. In addition, GhEREB2 and GhEREB3 proteins specifically bind to a GCC-box and strongly activate the expression of *HIS3* and *LacZ* reporter genes in yeast. Collectively, these results suggest that *GhEREB2* and *GhEREB3* might be involved as positive transcription factors in biotic stress signal transduction pathways.

DOI: 10.1134/S0006297906030084

Key words: AP2/EREBP domain, EREBP protein, GCC-box, GhEREB2, GhEREB3, cotton

Plants are often exposed to various biotic and abiotic stresses, such as a pathogen, which have adverse effects on the growth of plants. Once attacked by a pathogen, plants respond to the pathogen attack by expressing a large number of genes and synthesizing a large number of stress proteins that have putative roles in stress adaptation and plant defense [1, 2]. Transcriptional regulation of defense genes is thought to be important for the establishment of plant disease resistance [3]. Several plant transcription factors play roles in defense gene regulation, such as EREBP (ethylene response element binding protein) transcription factors, which is a crucial part of the plant response to pathogen stress by regulating the expression of PR (pathogenesis-related) genes [4-6].

EREBP proteins share a highly conserved AP2/EREBP domain of about 58-70 amino acids with no apparent similarity outside this DNA-binding domain [7]. This DNA-binding domain was first identified in APETALA2 (AP2) [8], and later found to be conserved in tobacco EREBPs, which are involved in plant defense [9]. To date, the AP2/EREBP domain has been found in various plant regulatory genes, which recognize and bind to some *cis*-acting elements, such as DRE (dehydration-responsive element), CRT (C-repeat) element, and GCC-box. The AP2/EREBP domain of EREBP proteins specifically bind with GCC-box (TAAGAGCCGCC), an ethylene-responsive promoter element found in many PR genes, which has been identified as an important element in biotic stress-responsive gene expression in an ethylene dependent manner [5, 10]. The core sequence (AGCCGCC) of the GCC-box is the optimal binding position, especially the 2nd G, 5th G, and 7th C of the GCC-box are involved in highly specific interactions with EREBP proteins [11]. The solution structure of the GCC-box binding domain consists of a three-stranded anti-parallel β-sheet and an α-helix packed approximately parallel to the β-sheet, and the β-sheet located in the N-terminal of AP2/EREBP domain has important roles in the binding with GCC-box [12]. Additionally, the α-helix and the

Abbreviations: AP2) APETALA2; 3-AT) 3-aminotriazole; CRT) C-repeat; DRE) dehydration-responsive element; EREBP) ethylene response element binding protein; JA) jasmonic acid; MCS) multiple cloning site; mGCC) mutant GCC; PR) pathogenesis-related; SA) salicylic acid.

* To whom correspondence should be addressed.

loop at the C-terminal of AP2/EREBP domain could play a role in protein–protein interactions to facilitate DNA binding, and then enlarge the target genes of EREBP proteins [13].

EREBP transcription factors not only play key roles in plant responses to various stresses and pathogen, but also are induced by various stresses, by hormones associated with defense responses, such as ethylene, jasmonic acid (JA), salicylic acid (SA), and by pathogens [9, 13-16]. With the isolation of the first EREBP proteins from tobacco [9], many EREBP transcription factors have been identified as GCC-box binding proteins in *Arabidopsis*, tomato, and other plants. In addition, Tsi1 and NtCEF1, two tobacco EREBP proteins, which could bind both the GCC-box and DRE element, have also been identified [14, 17].

In this study, *GhEREB2* and *GhEREB3* genes were isolated via yeast one-hybrid system from cotton and both had one single intron in their coding regions. *GhEREB2* was expressed in all of the tissues tested, but *GhEREB3* was only expressed in leaves. Furthermore, they both were induced by ethylene and JA, and *GhEREB2* also could be induced by drought and cold. In addition, their transcription activation ability was studied in yeast.

MATERIALS AND METHODS

Plant materials and stress treatments. Seeds of cotton (*Gossypium hirsutum* L., XH14), which were kindly provided by the Cotton Research Institute, Chinese Academy of Agricultural Sciences, were soaked at 24°C for 12 h, surface-sterilized for 30 sec with 75% ethanol, washed three times with sterile water, and then subsequently grown on Murashige and Skoog culture at 24°C under 16 h photoperiod. After the cotton plants were grown for two weeks, the plants were carefully pulled out and subjected to various stresses under dim light including dehydration, low temperature, ethylene, jasmonic acid (JA), and salicylic acid (SA) treatment. For dehydration (drought) treatment, the plants were carefully pulled out and placed on dry filter paper and allowed to dry at 30°C with a gentle air flow. The cold stress treatment was performed by transferring the plants to an incubator in which the temperature had been set at 4°C. For other stress conditions, the plants were transferred to solution containing 1 mM of ethylene, 100 μM of JA, or 2 mM of SA for ethylene, JA, and SA treatment, respectively, at 22°C. In each case, the plants were subjected to the stress treatments for various periods and then quickly immersed in liquid nitrogen and then stored at −80°C for DNA and RNA extraction.

DNA and RNA isolation. Total genomic DNA was extracted from cotton seedlings by a CTAB (cetyltriethylammonium bromide) extraction procedure as described previously [18] with modifications. Total RNA was extracted from cotton seedlings, roots, stems, and leaves with total RNA isolation system (Promega Corporation, USA) according to the instructions for product Z5110 and Z5651 at www.promega.com. The yields of genomic DNA and total RNA were determined by spectrophotometry at 260 nm, and the purity of genomic DNA and total RNA were estimated, also by spectrophotometry. The integrity of genomic DNA and total RNA were determined by denaturing agarose gel electrophoresis.

Construction of cDNA library. The poly(A+) mRNA was purified from total RNA of the cotton seedlings stressed with 1 mM of ethylene for 3 h with the PolyATract mRNA Isolation System (Promega Corporation), and 5 μg polyA mRNA were used for cDNA library construction with SuperScript™ Plasmid System (Invitrogen Corporation, USA) following the manufacturer's protocols. The cDNA library contained approximately $5 \cdot 10^6$ clones and was amplified. Recombinant cDNAs were excised and then cloned in MCS (multiple cloning sites) downstream of the GAL4 activation domain in pADGAL4 phagemid vector containing the *LEU*2 selective marker (Stratagene, USA).

Yeast one-hybrid screening. The yeast YM4271 strains carry two yeast expression vectors, p*HISi*-1 and p*Lac*Zi, which were constructed by Dr. Yuxia Hou and respectively contain reporter genes *HIS3* or *LacZ* fused to three copies of the GCC-box sequence (TAAGAGCCGCC) from the promoter region of PR, and can grow on SD/His-Ura selecting medium plates. Approximately 10^6 yeast transformants were screened using cDNA libraries prepared from ethylene-treated cotton seedlings according to the manufacturer's protocol (Clontech (USA) one-hybrid system). We obtained 15 positive colonies, which could grow on SD/His-Ura-Leu medium plates in the presence of 3-aminotriazole (3-AT) (a competitive inhibitor of the *HIS* gene product) and show β-galactosidase activity. The cDNAs of positive clones were excised from the pADGAL4 plasmids and ligated into the pBluescript II SK+ vector and then sequenced.

DNA sequencing and data analysis. The cDNA fragments were sequenced by GeneCore BioTechnologies Co. Ltd. (China) using the ABI373A automatic sequencer. DNA sequence data were analyzed using the DNASTAR and DNAMAN programs. The nucleotide and amino acid sequences were compared in the NCBI (National Center for Biotechnology Information) database using the BLAST analysis sever. Prediction of AP2/EREBP domain of the proteins was performed at Conserved Domain Search with the CDART tool (http://www.ncbi.nlm.nih.gov/Structure/Lexington/lexington.cgi/). Amino acid sequences and cDNA sequences were aligned with the DNAMAN program.

***Trans*-activation experiment with yeast.** To analyze *trans*-activation activity of isolated cDNA clones using the yeast one-hybrid system, the entire coding regions of cDNAs were prepared by PCR and cloned into the

*Eco*RI-*Sal*I site of MCS in Yep-GAP expression vector, a yeast expression vector without GAL4 domain. The recombinant Yep plasmids were transformed into the yeast YM4271 strain containing dual reporter genes *HIS3* and *LacZ* under the control of the promoter region of PR containing the GCC (TAAGAGCCGCC) or mutated GCC (TAAGAcCCgCg) for *trans*-activation experiments. The mutant GCC is also called mGCC. The transformed yeast cells (whether containing wild type GCC *cis*-element or mutant GCC *cis*-element) with two selective genes *HIS* and *LacZ* were streaked on the plates of SD/−His (without His) in the absence or presence of 3-AT and their growth was examined. The β-galactosidase activities of these colonies were analyzed according to the manufacturer's instructions (Clontech).

RT-PCR analysis. To study the expression of specific genes, 1 μg RNA isolated from cotton seedlings as described above was analyzed with One Step RNA PCR Kit (AMV) (TaKaRa Biotechnology Co., Ltd., Japan) according to the manufacturer's instruction. The RT-PCR primers of specific genes were designed according to the cDNA sequences. The constitutively expressed ubiquitin gene was used as the internal standard for RT-PCR. The primers of specific gene and ubiquitin gene designed for RT-PCR amplification were following: EF2, 5′-ATGTGTGGAGGTGCAATTA-3′; ER2, 5′-TCTGTTGTTGATGATGTGTC-3′; EF3, 5′-ATGTGTGGAGGTGCAATTA-3′; ER3, 5′-TAGTCTCATGGTTAGTCTCC-3′; UbF, 5′-CAGATCTTCGTCAAAACCCT-3′; UbR, 5′-GACTCCTTCTGGATGTTGTA-3′. The PCR products were separated by 1.2% agarose gel electrophoresis.

RESULTS

Isolation and structural analysis of two EREBP-like genes. To isolate cDNAs encoding EREBP transcription factors, we treated cotton seedlings with 1 mM ethylene for 3 h and constructed one cDNA library using SuperScript™ Plasmid System. Approximately 10^6 yeast transformants were screened according to the manufacturer's protocol (Clontech Mathchmaker one-hybrid system). Subsequently, 15 positive clones were identified and sequenced. Sequence analyses of these clones revealed that two cDNA clones have full open reading frame either with 1064 and 995 bp, respectively, and designated *GhEREB2* and *GhEREB3*. The *GhEREB2* and *GhEREB3* cDNA sequences have been submitted to the GenBank databases with GenBank accession numbers AY962571 and AY962572, respectively.

GhEREB2 contains a single open reading frame of 256 amino acids and encodes a putative EREBP protein with a predicted molecular mass of 28.8 kD (Fig. 1). *GhEREB3* contains a single open reading frame of 261 amino acids and encodes a putative EREBP protein with a predicted molecular mass of 29.6 kD (Fig. 1). The deduced amino acid sequences GhEREB2 and GhEREB3 showed 68.7% identities, and surprisingly, the DNA binding domains are closely related with 94.7% identities to each other. Database searches revealed that the amino acid sequence of GhEREB2 and GhEREB3 contained a conserved DNA binding domain of 58 amino acid residues, which showed obvious similarity to the AP2/EREBP domain of a large family of plant DNA-binding proteins (with at least 66.7% identity) (Fig. 2). Other typical features of transcription factors are present in the GhEREB2 and GhEREB3 proteins, such as the basic amino acid sequences in their N-terminal regions that might function as a nuclear localization signal, and the acidic regions that might act as an activation domain for transcription. All the results showed that GhEREB2 and GhEREB3 might carry out their presumed role as regulators of transcription in cotton.

In addition, the genomic DNA sequences of *GhEREB2* and *GhEREB3* were cloned by PCR with the two pairs of primers—EF2 and ER2, EF3 and ER3 (in "Materials and Methods")—which were designed according to the terminal cDNA sequences of their encoding regions, respectively. The specific PCR products were generated by using these primers, 1199 and 1296 bp, respectively, and cloned into pGEM-T easy vector (Promega) for sequencing. By comparing the genomic DNA and cDNA sequences of *GhEREB2* and *GhEREB3*, we found that the two genes each contain one single intron with 366 and 505 bp, respectively, in their coding regions.

Phylogenetic analysis of *GhEREB2*, *GhEREB3*, and *Arabidopsis* EREBP genes. Database searches with the putative amino acid sequences of GhEREB2 and GhEREB3 for homologous EREBP proteins from other plants were performed using the BLAST program, and we found they showed obvious similarity to some EREBP proteins which have known or implied roles in response to various stresses, such as, with 44-47% identity to ERF protein (BAC56862) from *Solanum tuberosum* and RAP2.3 (NP188299) from *Arabidopsis* (Fig. 3a). Furthermore, their DNA-binding domains are highly similar to each other (87.7-94.7% identity), and the N- and C-termini of these proteins also share significant similarity. In contrast, the amino acid sequences of GhEREB2 and GhEREB3 outside the above regions showed low significant sequence identity with *Arabidopsis* RAP2.3 and *S. tuberosum* ERF protein, which was also observed among other EREBP proteins [13, 14].

To further analyze the phylogenic relationships of GhEREB2, GhEREB3, and other EREBP from *Arabidopsis*, a gene tree was constructed based on the sequence identities of the AP2/EREBP domain. As shown in Fig. 3b, the EREBP proteins can be divided into three small subgroups, and GhEREB2 and GhEREB3 cluster with RAP2.3 with high bootstrap values. In addi-

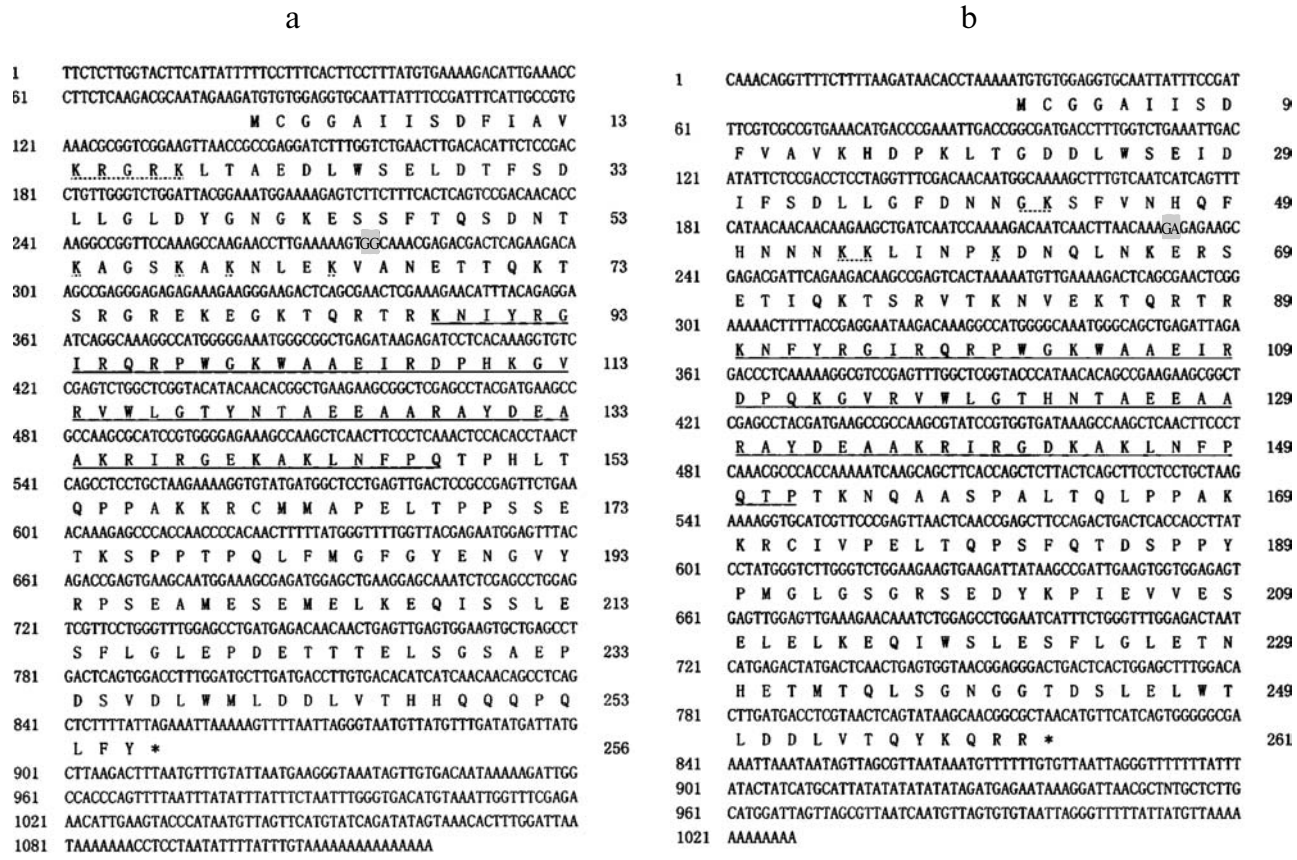

Fig. 1. Full-length cDNA and putative amino acid sequences of *GhEREB2* (a) and *GhEREB3* (b) genes. The amino acids that comprise the putative nuclear localization signal are shown by the dotted underline. The AP2/EREBP domain is indicated by the single underline. The nucleotide acids shadowed exhibit the position of intron of *GhEREB2* and *GhEREB3* genes.

Fig. 2. Alignment of AP2/EREBP domains of GhEREB2, GhEREB3, and other EREBP proteins. The AP2/EREBP domains of GhEREB2, GhEREB3, and other EREBP, such as AtERF1-3, RAP2.1-2.3 in *Arabidopsis thaliana*, EREBP1-3 in *Nicotiana tabacum*, and Pti4-6 in *Lycopersicon esculentum*, are compared by the DNAMAN software. The identity of amino acid residues is shown dark. The consensus indicates the conserved amino acids in AP2/EREBP domains of these EREBP proteins.

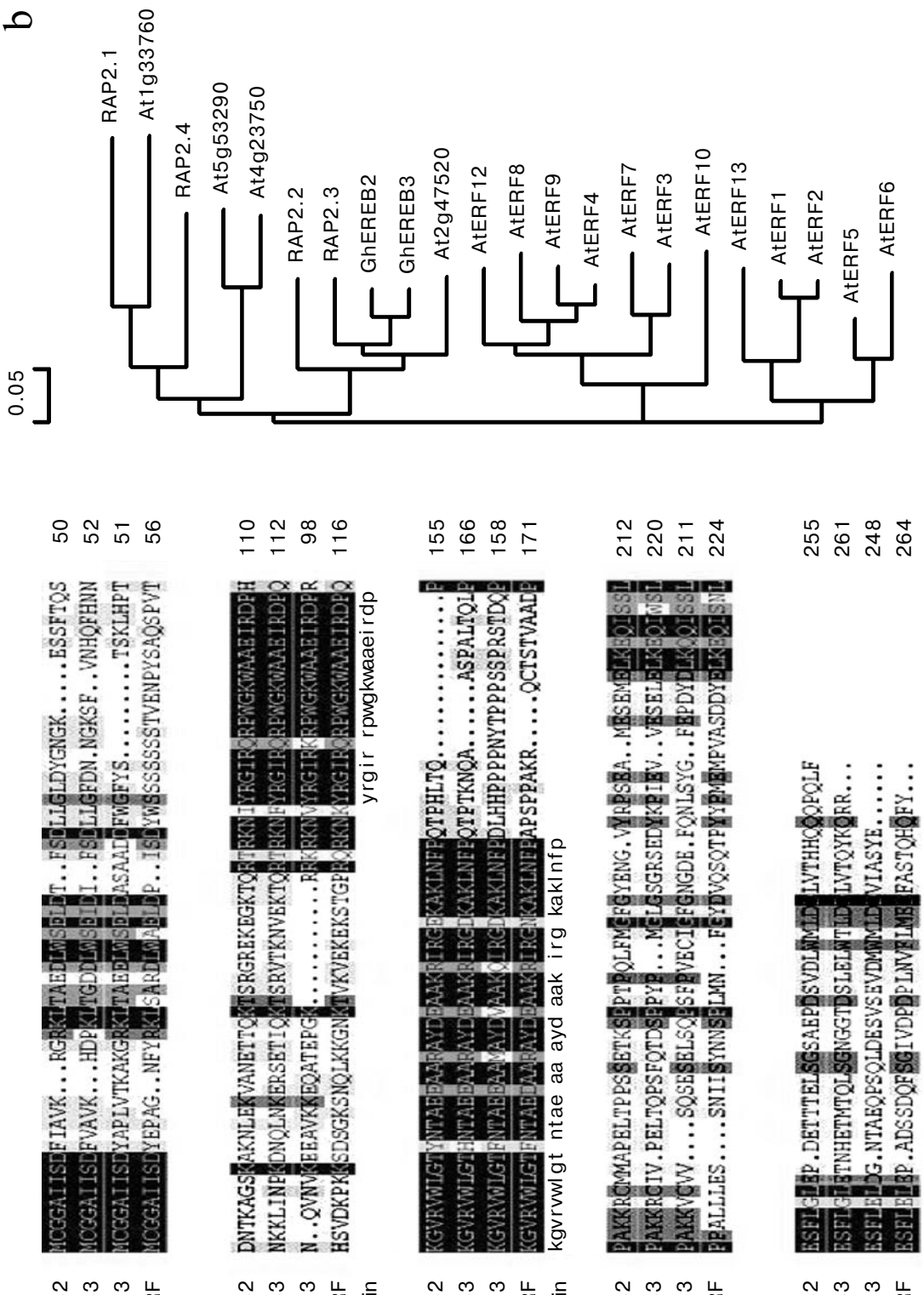

Fig. 3. Comparison of putative amino acid sequences of GhEREB2 and GhEREB3 with other EREBP proteins. a) Sequence alignment of GhEREB2, GhEREB3, and other EREBP proteins. Putative amino acid sequences of GhEREB2, GhEREB3, *Arabidopsis* RAP2.3 (NP188299), and *Solanum tuberosum* ERF protein (BAC56862) were aligned by the DNAMAN software. Identity and similarity of these three proteins are shown dark or gray, respectively. b) Phylogenic tree of AP2/EREBP domains of GhEREB2, GhEREB3, and EREBP proteins from *Arabidopsis*. The AP2/EREBP domains of GhEREB2 and GhEREB3 are compared with RAP2.1 (NP564496), RAP2.2 (NP850582), RAP2.3 (AAM65031), RAP2.4 (AAC49770), RAP2.6 (AAC36019), At1g33760 (AAV31158), At2g47520 (AAL69461), At4g23750 (AAT70489), At5g53290 (AAV74238), AtERF1 (QO80337), AtERF2 (AAM64544), AtERF3 (NP175479), AtERF4 (AAM64308), AtERF5 (BAA97157), AtERF6 (Q8VZ91), AtERF7 (Q9LDE4), AtERF8 (Q9MAI5), AtERF9 (BAB18560), AtERF10 (Q9ZWA2), AtERF11 (Q9C5I3), AtERF12 (NP174158), and AtERF13 (Q8L9K1) by the DNAMAN software. The scale bar indicates the branch length.

tion, they also show higher bootstrap values with RAP2.2 and At2g4750, and moderately high bootstrap values with RAP2.1, RAP2.4, At1g33760, At4g23750, and At5g53290, but are more distantly related to AtERF 1-13. Thus, GhEREB2 and GhEREB3 belong to the B3 subgroup of EREBP proteins basing on Sakuma et al. [19], and the high degree of homology shared between GhEREB2, GhEREB3, and RAP2.3 indicates that RAP2.3 may be the functional similarity of GhEREB2 and GhEREB3.

Expression pattern of *GhEREB2* and *GhEREB3*. Since Northern blot analysis produced only weak or undetectable signals (results not shown), we analyzed the expression of *GhEREB2* and *GhEREB3* genes in a variety of organs of cotton and under various stress conditions by RT-PCR. Expression of *GhEREB2* and *GhEREB3* genes are detected in the cotton plants, which were untreated, and shown tissue organ-specific expression pattern. The level of *GhEREB2* expression is the highest in the stems, higher in the roots, and lower in the leaves (Fig. 4a). Surprisingly, the *GhEREB3* expression was only detected in leaves with a higher expression level than *GhEREB2*.

Ethylene has been shown to play important roles in a number of plant stress responses and in the expression of some EREBP genes [10, 20]. *GhEREB2* and *GhEREB3* were strongly induced by ethylene, their transcript levels increased markedly within 1 h and reached maximum at 9 h, and then their transcript levels began to fall (Fig. 4b). Like ethylene, SA and JA are important phytohormones involved in signaling in response to pathogen infection, and as shown in Fig. 4b, *GhEREB2* and *GhEREB3* were both upregulated by JA treatments, but did not display observable increase in response to SA (results not shown). Furthermore, the levels of *GhEREB2* and *GhEREB3* expression were higher under ethylene treatment than the JA treatment. Interestingly, *GhEREB2* was also slightly induced by stress conditions, such as drought and cold treatment, the transcript level of the *GhEREB2* treated after 3 h being only slightly higher than those found in control seedlings and peaking at 7 h. However, GhEREB3 transcripts did not accumulate in response to drought and cold treatments (Fig. 4b), indicating that *GhEREB3* was specifically regulated by ethylene and JA. Taken together, these results indicate that *GhEREB2* and *GhEREB3* have important roles in ethylene and JA signal transduction pathways, and furthermore, *GhEREB2* may be involved in cold- and drought-stress signal transduction pathways.

Transcriptional regulation of *GhEREB2* and *GhEREB3*. In order to further confirm the *trans*-activation ability of *GhEREB2* and *GhEREB3* genes, *trans*-activation experiments were performed in yeast. We ligated the positive cDNA fragments to the yeast expression vector Yep-GAP without GAL4 domain and transformed into yeast cells carrying *HIS3* and *LacZ* dual reporter genes fused to three copies of the wild-type or mutant GCC-box sequence (Fig. 5a). As shown in the table, all of the transformants can grow well on the selective medium without histidine and 3-AT, and in wild-type GCC yeast, transformants carrying the *GhEREB2* or *GhEREB3* plas-

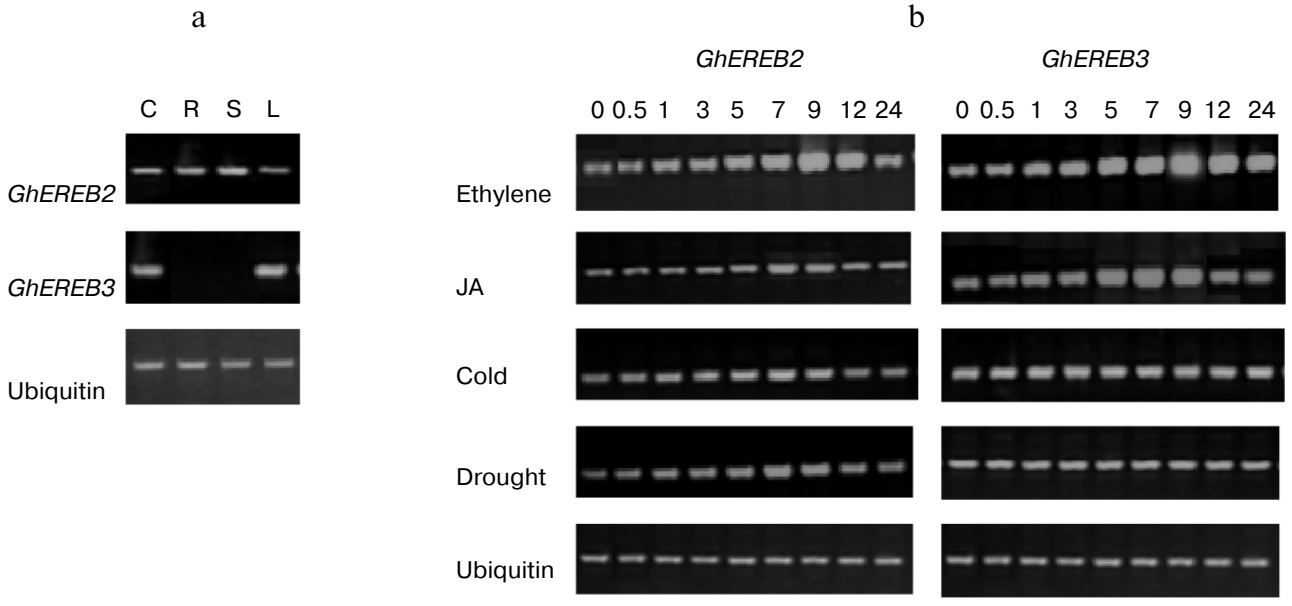

Fig. 4. RT-PCR analysis of expression pattern of *GhEREB2* and *GhEREB3* genes. a) Expression of the *GhEREB2* and *GhEREB3* genes in the organs of normally grown cotton. The capital letters C, R, S, and L, respectively, designate whole plants, roots, stems, and leaves of cotton plants. b) Expression of the *GhEREB2* and *GhEREB3* genes in response to ethylene, SA, JA, low temperature, or dehydration. The number above each lane indicates the number of hours after the initiation of treatment before the isolation of total RNA. In addition, the internal controls (ubiquitin) are identical in four different experiments.

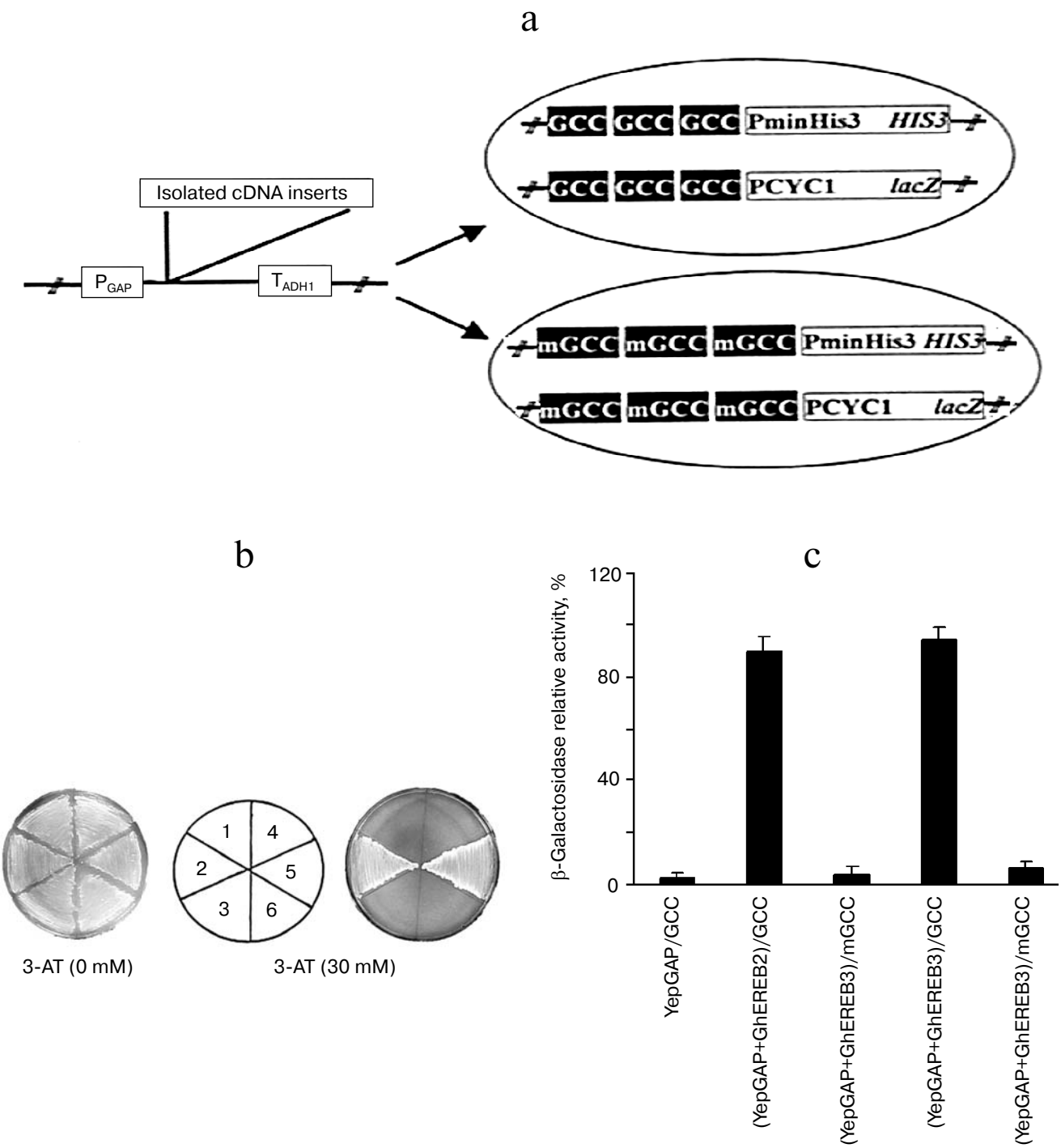

Fig. 5. Activation of *His* and *LacZ* reporter genes by GhEREB2 and GhEREB3 in yeasts. a) The cDNA fragments of *GhEREB2* and *GhEREB3* were cloned into the yeast expression vector YepGAP and transformed into yeast carrying dual reporter genes *HIS3* and *LacZ* under the control of the promoter region containing GCC or mGCC. P_{GAP} indicates the promoter of glyceraldehyde 3-phosphate dehydrogenase gene, and T_{ADH1} indicates the terminator of the *ADH1* gene. b) The plasmid containing *GhEREB2* or *GhEREB3* was transformed into yeast carrying dual reporter genes containing GCC-box or mGCC-box. The transformants were examined for growth in the presence of 30 mM 3-AT. *1*) YepGAP/GCC; *2*) (YepGAP + GhEREB2)/GCC; *3*) (YepGAP + GhEREB2)/mGCC; *4*) YepGAP/GCC; *5*) (YepGAP + GhEREB3)/GCC; *6*) (YepGAP + GhEREB3)/mGCC. c) *LacZ* expression in yeast cells transformed with *GhEREB2* or *GhEREB3*. The yeast cells were grown on a selective medium plate of SD/His– for 22 h. The β-galactosidase values are from at least three independent replicates. Error bars represent standard errors.

Activation of *HIS* reporter gene in yeast cells transformed with *GhEREB2* or *GhEREB3*

Reporter strain	3-AT concentration, mM						
	0	10	20	30	40	50	60
(YepGAP + GhEREB2)/GCC	+	+	+	+	+	+	+
(YepGAP + GhEREB2)/mGCC	+	—	—	—	—	—	—
(YepGAP + GhEREB3)/GCC	+	+	+	+	+	+	+
(YepGAP + GhEREB3)/mGCC	+	—	—	—	—	—	—
YepGAP/GCC	+	—	—	—	—	—	—

Note: The transformants were examined for growth on selective medium plate of SD/His– with 3-AT at indicated concentration. Growth of colonies was evaluated after six days: –, no growth; +, good growth.

mid can grow well on the selective medium lacking histidine in the presence of 10, 20, 30, 40, 50, and 60 mM 3-AT (Fig. 5b), whereas the transformants carrying mGCC-box or the negative control cells did not show 3-AT resistance. These results indicated that GhEREB2 and GhEREB3 can bind to the GCC-box and activate the expression of the reporter gene *HIS3*.

The quantitative assay of β-galactosidase activity showed that the intensities of *LacZ* expression increased when the report plasmid containing wild-type GCC-box sequence was co-expressed with the *GhEREB2* or *GhEREB3* effector plasmid. On the contrary, compared to the negative control, there was no significant change of *LacZ* expression in the yeast cells containing mGCC-box sequence (Fig. 5c). Furthermore, the β-galactosidase activity in the yeast containing *GhEREB3* was corresponding higher than that in the yeast containing *GhEREB2*. Therefore, the results showed that heterogeneous expression of GhEREB2 and GhEREB3 both promoted the expression of *HIS3* and *LacZ* genes in yeast with wild-type-box but not in yeast with mutant GCC-box, indicating that GhEREB2 and GhEREB3 encoding products could specifically bind to GCC-box and activate target gene expression *in vivo*.

DISCUSSION

EREBP proteins belong to a superfamily of plant transcription factors and have important regulatory functions in plant development, environment stress responses, and defense response [4, 6, 14, 16, 21, 22]. In this research, we isolated two cDNA clones from cotton cDNA library by yeast one-hybrid system, designated *GhEREB2* and *GhEREB3*, which each has two exons and one intron in their coding regions. GhEREB2 and GhEREB3 proteins belong to new members of the EREBP transcription factor family and each contains one AP2/EREBP domain, a nuclear localization signal, and an activation region. The deduced amino acid sequences of GhEREB2 and GhEREB3 have higher similarity, especially with 97.4% similarity in the DNA-binding domain. Database searches showed that the two proteins share obvious similarity to ERF protein (BAC56862) from *S. tuberosum* and RAP2.3 (NP188299) from *Arabidopsis*, but they showed low significant sequence similarity except for the DNA-binding domain and N- and C-terminus regions, which was also observed among other EREBP proteins [13, 14]. Additionally, we analyzed the phylogenic relationships of GhEREB2, GhEREB3, and some EREBP proteins from *Arabidopsis* basing on the AP2/EREBP domain and found that the two proteins may belong to the B3 subgroup of EREBP proteins according to Sakuma et al. [19], and RAP2.3 may be one functional homolog of GhEREB2 and GhEREB3.

The expression of *GhEREB2* and *GhEREB3* were constitutive and affected by some stresses. *GhEREB2* was expressed in roots, stems, and leaves of the cotton, indicating that *GhEREB2* may also function in the normal program of plant growth and development. In contrast, *GhEREB3* was only expressed in the leaves, but the level of *GhEREB3* was higher than the *GhEREB2*. We also found interesting differences in the expression of *GhEREB2* and *GhEREB3* genes in response to defense signaling molecules, such as ethylene, SA, and JA, which are important components of defense response pathways [2, 23, 24]. *GhEREB2* and *GhEREB3* were both induced by ethylene and JA, thus they may be involved in the ethylene/JA-dependent signaling pathway regulating the expression of vacuole localized basic PR proteins such as PR3, PR4, PR5, and PDF1.2, as the AtERF1 from *Arabidopsis* [25]. In addition, *GhEREB2* was slightly upregulated by drought and cold conditions, indicating *GhEREB2* is also likely to be regulated by signaling pathways under cold and drought stress.

EREBP proteins have been shown to function as either transcriptional activators or repressors [10, 26], such as AtERF7 binds to the GCC-box and acts as a

repressor of gene transcription [16]. The GhEREB2 and GhEREB3 proteins both contain a highly conserved AP2/EREBP DNA-binding domain and an acid-rich region that may function as transcriptional activation domains. In this study, using a yeast one-hybrid system, we demonstrated the *in vivo* transcriptional properties of GhEREB2 and GhEREB3, and found that they all both promoted expression of the *HIS3* and *LacZ* genes in wild-type GCC yeast but not in mutant GCC yeast, indicating that GhEREB2 and GhEREB3, similarly to EREBP proteins, could specifically recognize and bind to GCC-box motif that is present in the promoter region of a large number of genes encoding PR proteins and activate target gene expression *in vivo*. These data suggested that GhEREB2 and GhEREB3 proteins might function as a transcriptional regulator involved in the GCC-dependent gene expression in cotton.

Taken together, we have identified *GhEREB2* and *GhEREB3* genes from cotton, and their encoding proteins have a highly conserved AP2/EREBP domain and other typical features of transcription factors. The expression patterns of *GhEREB2* and *GhEREB3* are different, which may reflect on their different function. Given the high similarity between GhEREB2 and GhEREB3, it is not clear whether GhEREB2 and GhEREB3 are functionally redundant, and whether GhEREB2 might regulate gene expression in response to biotic stress in cotton. Therefore, further studies would reveal the function and molecular mechanism of GhEREB2 and GhEREB3.

We wish to thank the Cotton Research Institute, Chinese Academy of Agricultural Sciences for the supply of cotton seeds and Dr. Xianguo Cheng for helpful advice in the experiments.

This research was supported by the Genetic Transformation on Plants Project (JY2002-B-001) and 863 Project (2004AA212101), P. R. China.

REFERENCES

1. Skriver, K., and Mundy, J. (1990) *Plant Cell*, **2**, 502-512.
2. Reymond, P., and Farmer, E. E. (1998) *Curr. Opin. Plant Biol.*, **1**, 404-411.
3. Lamb, C. J., Lawton, M. A., Dron, M., and Dixon, R. A. (1989) *Cell*, **56**, 215-224.
4. Zhou, J., Tang, X., and Martin, G. (1997) *EMBO J.*, **16**, 3207-3232.
5. Kitajima, S., and Sato, F. (1999) *J. Biochem.*, **125**, 1-8.
6. Guo, Z. J., Chen, X. J., Wu, X. L., Ling, J. Q., and Xu, P. (2004) *Plant Mol. Biol.*, **4**, 607-618.
7. Okamuro, J. K., Caster, B., Villarroel, R., Montagu, M., and Jofuku, K. D. (1997) *Proc. Natl. Acad. Sci. USA*, **94**, 7076-7081.
8. Jofuku, K. D., den Boer, B. G. W., van Montagu, M., and Okamura, J. K. (1994) *Plant Cell*, **6**, 1211-1225.
9. Ohme-Takagi, M., and Shinshi, H. (1995) *Plant Cell*, **7**, 173-182.
10. Fujimoto, S. Y., Ohta, M., Usui, A., Shinshi, H., and Ohme-Takagi, M. (2000) *Plant Cell*, **12**, 393-404.
11. Hao, D. Y., Ohme-Takagi, M., and Sarai, A. (1998) *J. Biol. Chem.*, **273**, 26857-26861.
12. Allen, M. D., Yamasaki, K., and Ohme-Takagi, M. (1998) *EMBO J.*, **18**, 5484-5496.
13. Buttner, M., and Singh, K. B. (1997) *Proc. Natl. Acad. Sci. USA*, **94**, 5961-5966.
14. Park, J. M., Park, C. J., Lee, S. B., Ham, B. K., Shin, R., and Paek, K. H. (2001) *Plant Cell*, **13**, 1035-1046.
15. Shen, H., Yang, H. J., and Wang, Z. Y. (2004) *Zhi Wu Sheng Li Yu Fen Zi Sheng Wu Xue Xue Bao*, **1**, 87-93.
16. Song, C. P., Agarwal, M., Ohta, M., Guo, Y., Halfter, U., Wang, P., and Zhu, J. K. (2005) *Plant Cell*, **8**, 2384-2396.
17. Jae, H. L., Do, M. K., Jeong, H. L., Jungmook, K., Jae, W. B., Woo, T. K., and Hyun, S. P. (2005) *Planta* (E-pub ahead of print) PMID: 15918028.
18. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1987) *Current Protocols in Molecular Biology*, John Wiley, New York.
19. Sakuma, Y., Liu, Q., Dubouzet, J. G., Abe, H., Shinozaki, K., and Shinozaki, K. Y. (2002) *Biochem. Biophys. Res. Commun.*, **290**, 998-1009.
20. Solano, R., Stepanova, A., Chao, Q., and Ecker, J. R. (1998) *Genes Dev.*, **12**, 3703-3714.
21. Yang, H. J., Shen, H., Chen, L., Xing, Y. Y., Wang, Z. Y., Zhang, J. L., and Hong, M. M. (2002) *Plant Mol. Biol.*, **50**, 379-391.
22. Aharoni, A., Dixit, S., Jetter, R., Thoenes, E., van Arkel, G., and Pereira, A. (2004) *Plant Cell*, **9**, 2463-2480.
23. Dong, X. (1998) *Curr. Opin. Plant Biol.*, **1**, 316-323.
24. Dempsey, D. A., Shah, J., and Klessig, D. F. (1999) *Plant Sci.*, **18**, 547-575.
25. Lorenzo, O., Piqueras, R., Sanchez-Serrano, J. J.., and Solano, R. (2003) *Plant Cell*, **15**, 165-178.
26. Ohta, M., Matsui, K., Hiratsu, K., Shinshi, H., and Ohme-Takagi, M. (2001) *Plant Cell*, **13**, 1959-1968.