RESEARCH PAPER

Phylogeny, gene structures, and expression patterns of the ERF gene family in soybean (Glycine max L.)

Gaiyun Zhang1,2,*, Ming Chen2,*, Xueping Chen1,†, Zhaoshi Xu2, Shan Guan2, Lian-Cheng Li2, Aili Li2, Jiaming Guo1, Long Mao2 and Youzhi Ma2,†

1 Department of Chemistry, University of Science and Technology of China, Hefei 230026, China
2 The National Key Facility for Crop Genetic Resources and Genetic Improvement, Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing, 100081, China

Received 15 May 2008; Revised 7 September 2008; Accepted 11 September 2008

Abstract

Members of the ERF transcription factor family play important roles in regulating gene expression in response to biotic and abiotic stresses. In soybean (Glycine max L.), however, only a few ERF genes have been studied so far. In this study, 98 unigenes that contained a complete AP2/ERF domain were identified from 63 676 unique sequences in the DFCI Soybean Gene Index database. The phylogeny, gene structures, and putative conserved motifs in soybean ERF proteins were analysed, and compared with those of Arabidopsis and rice. The members of the soybean ERF family were divided into 12 subgroups, similar to the case for Arabidopsis. AP2/ERF domains were conserved among soybean, Arabidopsis, and rice. Outside the AP2/ERF domain, many soybean-specific conserved motifs were detected. Expression analysis showed that nine unigenes belonging to six ERF family subgroups were induced by both biotic/abiotic stresses and hormone treatment, suggesting that they were involved in cross-talk between biotic and abiotic stress-responsive signalling pathways. Overexpression of two full-length genes from two different subgroups enhanced the tolerances to drought, salt stresses, and/or pathogen infection of the tobacco plants. These results will be useful for elucidating ERF gene-associated stress response signalling pathways in soybean.

Key words: ERF family, gene function, phylogeny, soybean, stress response.

Introduction

Drought, high salinity, low temperature, and pathogen attack are the most common stress factors that influence plant growth and development. To overcome these limitations, plants respond and adapt to stresses at the physiological and biochemical levels. AP2/ERF transcription factors, characterized by the presence of the AP2/ERF DNA-binding domain, play significant roles in regulating plant biotic and abiotic stress-responsive gene expression, (Sakuma et al., 2002). AP2/ERF genes constitute a large superfamily, which has been divided into three groups named the AP2, ERF, and RAV families based on their sequence similarities and numbers of AP2/ERF domains (Nakano et al., 2006). AP2 proteins contain two AP2/ERF domains, and genes in this family participate in the regulation of developmental processes (Elliott et al., 1996; Chuck et al., 1998; Boutilier et al., 2002). RAV family proteins contain one AP2/ERF domain and a B3 domain, and have different biological functions compared with members in other families. Recently, members of the RAV family were shown to be involved in the ethylene response (Alonso et al., 2003), the brassinosteroid response (Hu et al., 2004), and biotic and abiotic stress responses (Sohn et al., 2006). ERF family proteins contain a single AP2/ERF domain, and are sometimes further divided into two major subfamilies, the CBF/DREB subfamily and the ERF subfamily (Sakuma et al., 2002). Genes in the CBF/DREB subfamily play a crucial role in the response of plants to abiotic stresses by recognizing the dehydration-responsive element (DRE) with a core motif of A/GCCGAC (Yamaguchi-Shinozaki and Shinozaki, 1994; Thomashow, 1999). The ERF subfamily genes are

* These authors contributed equally to this work.
† To whom correspondence should be addressed. E-mail: mayzh@mail.caas.net.cn

© 2008 The Author(s).
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/2.0/uk/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
mainly involved in response to biotic stresses such as pathogenesis by recognizing the cis-acting element AGCCGC, known as the GCC box (Hao et al., 1998).

ERF and CBF/DREB subfamily transcription factors have been identified in various plant species, including Arabidopsis (Liu et al., 1998; Ohate-Sánchez and Singh, 2002), rice (Cao et al., 2006), and cotton (Huang et al., 2007; Jin and Liu, 2008). The roles of ERF and CBF/DREB proteins in the response to biotic and abiotic stress have also been extensively documented (Gutteron and Reuber, 2004; Agarwal et al., 2006). The sequenced Arabidopsis genome contains 145 distinct genes encoding AP2/ERF-type proteins classified into five groups, that is APETALA2 (AP2) subfamily (17 genes), RAV subfamily (six genes), CBF/DREB subfamily (56 genes), ERF subfamily (65 genes), and one very specific gene, AL079349, based on similarities of their AP2/ERF DNA-binding domains (Sakuma et al., 2002). The proteins of the CBF/DREB subfamily were further divided into six subgroups, A-1 to A-6, among which A-1 and A-2 were the two largest (Sakuma et al., 2002). Expression of the DREB1A/CBF3 (A-1) genes is induced by low temperature stress, but not by drought or high salt stresses, whereas DREB2A (A-2) genes are induced by drought and high salt, but not by low temperature (Liu et al., 1998). Overexpression of DREB1A/CFB3 under control of the cauliflower mosaic virus (CaMV) 35S promoter increased tolerance to drought, high salt, and freezing stresses (Liu et al., 1998; Kasuga et al., 1999, Gilmour et al., 2000).

Overexpression of constitutively active DREB2A resulted in significant drought stress tolerance, but only slight freezing tolerance in transgenic Arabidopsis plants (Sakuma et al., 2006). Other DREB proteins such as TINY2 (A-4), GhDBP1 (A-5), GmDREB2 (A-5), and ZmDBF1 (A-6) were also characterized as stress-inducible proteins (Kizis and Pages, 2002; Wei et al., 2005; Huang and Liu, 2006; Chen et al., 2007). The proteins of the ERF subfamily were also divided into six groups termed B-1 to B-6. The expression and biological functions of genes in the ERF subfamily were summarized by Nakano et al. (2006). An example, transcription of tobacco Tsl (for Tobacco stress-induced gene1) was induced by salt, ethephon (ET), and salicylic acid (SA). Overexpression of Tsl improved the tolerance to salt and pathogen attack (Park et al., 2001). Nakano et al. (2006) reported 147 and 157 genes in Arabidopsis and rice, respectively, which were classified as members of the AP2/ERF superfamily. Among the Arabidopsis ERF genes, 122 were considered as members of the CBF/DREB subfamily and the ERF subfamily (Sakuma et al., 2002). In rice, there are 139 genes in the ERF family (Nakano et al., 2006). The phylogeny, gene structures, and conserved motifs of ERF genes in these two species were also analysed (Nakano et al., 2006).

Soybean is one of the most economically important crop species in the world. Only a few members of the ERF and CBF/DREB subfamily have been characterized in this species (Mazarei et al., 2002; Li et al., 2005), and most of their functions remain to be determined. Recently, four CBF/DREB homologous genes (GmDREBa/b/c and GmDREB2), and one ERF homologous gene (GmEREBP1) were isolated from soybean and their expression characteristics were investigated (Mazarei et al., 2002; Li et al., 2005; Chen et al., 2007). To gain further information about the AP2/ERF superfamily in soybean, the DFCI Soybean Gene Index database was surveyed and 148 unigenes were identified in this superfamily, including 120 ERF family unigenes, 26 AP2 family unigenes, and two RAV family unigenes. Phylogenetic and protein motif structural analyses of the ERF and CBF/DREB subfamily were undertaken. The expression patterns of some genes belonging to the ERF subfamily were also characterized. The biological functions of the two full-length ERF genes were investigated in transgenic tobacco plants. The results from this study, reported herein, form a basis for future functional analyses of the soybean ERF family genes.

Materials and methods

Data set and data treatment

The soybean unigene set and the expressed sequence tags (ESTs) used as a primary sequence data set are available on the TIGR Web site (http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/gimain.pl?gudb=soybean) as the DFCI Soybean Gene Index, from which 63,676 unique sequences weredownloadincluding31928TCstentativa consensus sequences), 31,636 singleton ESTs, and 112 singleton expressed transcript sequences. These data were released on 20 September 2004.

The Transeq program from the EMBOSS package was used to translate DNA sequences into protein sequences. The amino acid sequences of the longest open reading frame (ORF) in six ORFs were selected, and amino acid sequences whose lengths <100 were excluded for the following analysis. Based on the HMMER User’s Guide (http://hmmer.wustl.edu/ Version 2.3.2; Oct 2003), the Hmmpfam program was then used to annotate various kinds of domains in the query sequence, then the hmmpfam program was used to retrieve an HMM as a seed model from an HMM database, including the AP2/ERF domain. Finally, the hmmalign program was used to align multiple TC/EST sequences to the seed profile HMM, and 148 TC/EST sequences containing the AP2/ERF domain were obtained.

The Arabidopsis gene set is available through the Arabidopsis Information Resource (http://www.Arabidopsis.org). The rice genes of the ERF family were downloaded from the National Center for Biotechnology Information (NCBI) and TIGR rice genome annotation databases (http://rice.plantbiology.msu.edu/).

Alignment, phylogenetic analysis, and motif detection

All similarity searches were executed locally using the BlastN, BlastX, or BlastP tools at the NCBI, TIGR, and TAIR web sites. Conserved domain searches were performed against the conserved domain database at NCBI using the reversed position-specific-blast algorithm with translated unigenes. A phylogenetic tree was constructed with the aligned soybean AP2/ERF protein sequences
using MEGA (version 4.0; http://www.megasoftware.net; Tamura et al., 2007) and the Neighbor-Joining (NJ) method with the following parameters: Poisson correction, pairwise deletion, and bootstrap (1000 replicates; random seed). The amino acid variation rates were also obtained. Motif detection was performed with MEME (Bailey et al., 2006) (MIME version 3.5.7, http://meme.sdsc.edu/meme/meme.html).

Table 1. Summary of the structure of the AP2/ERF superfamily in soybean compared with Arabidopsis

<table>
<thead>
<tr>
<th>Group</th>
<th>Conserved domain</th>
<th>Soybean</th>
<th>Arabidopsis&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP2 family</td>
<td>Double complete or incomplete AP2/ERF domain</td>
<td>26</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Single complete or incomplete AP2/ERF domain</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>RAV family</td>
<td>Single AP2/ERF domain and one B3 domain</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>ERF family</td>
<td>Single AP2/ERF domain</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>CBF/DREB subfamily</td>
<td>Single AP2/ERF domain</td>
<td>22</td>
<td>121</td>
</tr>
<tr>
<td>ERF subfamily</td>
<td>Single complete AP2/ERF domain</td>
<td>36</td>
<td>56</td>
</tr>
<tr>
<td>AL070349</td>
<td>Single complete AP2/ERF domain</td>
<td>62</td>
<td>65</td>
</tr>
<tr>
<td>AP2/ERF superfamily</td>
<td>Total</td>
<td>148</td>
<td>145</td>
</tr>
</tbody>
</table>

<sup>a</sup> From Sakuma et al. (2002).
Phylogenetic relationships between members of the ERF family in soybean and Arabidopsis

To confirm classifications and to analyse the phylogenetic relationships, multiple alignment analyses of the amino acid sequences of the AP2/ERF domain in the 98 soybean ERF proteins were performed using the 122 Arabidopsis ERF proteins described by Nakano et al. (2006) (Supplementary Fig. S1 at JXB online) for comparison. Residues Glu17, Trp36, Leu37, and Gly38 were completely conserved among all 220 proteins in both species (Supplementary Fig. S1). In addition, >95% of the ERF family members contain Gly4, Arg6, Arg8, Gly12, Ile18, Arg19, Arg34, Ala46, Ala47, Asp51, and Asn65 residues. Based on alignment, an NJ phylogenetic tree was generated with bootstrap analysis (1000 replicates). As shown in Fig. 1 and Supplementary Fig. S2 at JXB online, the phylogenetic tree divided the ERF family proteins of Arabidopsis and soybean into 12 subgroups, designated A-1 to A-6 and B-1 to B-6, in accordance with the classification described by Sakuma et al. (2002). For example, CBF2/DREB1C (At4g25470) and DDF1/DREB1E (At1g63030) in subgroup A-1 in Arabidopsis (Sakuma et al., 2002) was also placed in subgroup A-1 in the present study. Subgroups A-1 to A-6 represent the CBF/DREB subfamily, and subgroups B-1 to B-6, the ERF subfamily. Comparative analyses of the phylogenetic tree suggested that the classification of the soybean ERF family was similar and applicable to the Arabidopsis ERF family.

Conserved motifs outside of the AP2/ERF domain in soybean and Arabidopsis

The conserved motifs in ERF family proteins in both soybean and Arabidopsis were investigated using MEME version 3.5.4, and the results for both species are listed in Supplementary Table S4 at JXB online. Most members in the same group shared one or more motifs outside the AP2/ERF domain (Figs. 2, 3). For example, the A-1 subgroup consisted of three unigenes (GmERF001 to 003) and contained three conserved motifs (Fig. 2). All of the unigenes in this subgroup contain the CMA-1-2 motif in the C-terminal region; this was reported as an LWSY conserved motif in OsDREB1A/B/C and in AtCBF3/DREB1A (Dubouzet et al., 2003). In addition, this is the CMA-1-1 motif located on both sides of the AP2/ERF domain in the proteins of GmERF002 and GmERF003 (Fig. 2).

Some conserved motifs identified in the Arabidopsis ERF family were also examined in the deduced amino acid sequences of GmERF unigenes. For example, the ERF-associated amphipathic repression (EAR) motif was found in members of subgroup A-5 as the CMA-5-2 motif, and in subgroup B-1 as the CMB-1-2 motif in both Arabidopsis and soybean (Supplementary Table S4 at JXB online, Figs 2–4). In addition, alanine-rich (in subgroup A-4 as the CMA-4-1 motif), glutamine-rich (in subgroup A-6 as the CMA-6-3 motif), and serine-rich (in subgroup A-4 as the CMA-4-7 motif, and in subgroup B-1 as the CMB-1-1 and CMB-1-3 motifs) amino acid sequences were detected in both Arabidopsis and soybean (Figs. 2, 3, Supplementary Table S4). These previously were often designated as transcriptional activation domains (Liu et al., 1999), but their functions were not rigorously demonstrated. An MCGRALL motif, designated as CMB-1-1, was a characteristic feature of subgroup B-1 in both species (Fig. 3, Supplementary Table S4).

In addition to the conserved motifs between soybean and Arabidopsis, there were also some soybean-specific motifs in the ERF family (Fig. 5, Supplementary Table S4). For example, the CMA-5-5 motif in subgroup A-5, the CMA-6-10 and CMA-6-11 motifs in subgroup A-6, the CMB-1-4, CMB-1-8, and CMB-1-9 motifs in subgroup B-1, the CMB-2-2 and CMB-2-8 motifs in subgroup B-2, and the CMB-3-9 motif in subgroup B-3 occurred only in soybean ERF proteins, but the functions of these motifs remain unknown.

Comparative analysis of the ERF gene family between soybean and rice

A total of 139 rice ERF family members were downloaded from the NCBI and TIGR rice genome annotation
To determine the phylogenetic relationships of the ERF family genes in soybean and rice, a multiple sequence alignment was performed using amino acid sequences in the AP2/ERF domain. This analysis revealed that those amino acid residues which might be involved in some form of physical contact with DNA are also conserved among most of the soybean ERF proteins and rice ERF proteins (Supplementary Fig. S3 at JXB online). Residues Arg8 and Gly39 were completely conserved among all 237 proteins in both species (Supplementary Fig. S3). In addition, >95% of the ERF family members contain Gly4, Arg6, Arg23, Arg33, Trp37, Leu38, Ala48, Ala50, Asp52, Asn75, and Phe76 residues (Supplementary Fig. S3). The phylogenetic tree containing soybean and rice ERF genes was constructed, and the phylogram were classed into 15 groups, namely groups I–XIV and a solo group VI-L (Supplementary Fig. S4 at JXB online), according to the report of Nakano et al. (2006). Among these groups, groups I–X were relevant to A1–A6 and B1–B6 shown in Supplementary Fig. S4 and Table S5 at JXB online. No soybean ERFs were assigned to groups XI–XIV and VI-L (Supplementary Fig. S4) that were specific to rice. Similarly, no Arabidopsis ERFs fell in the groups XI–XIV either (Nakano et al., 2006).

The comparative analysis of conserved motifs indicated that most of the motifs conserved in the soybean and Arabidopsis ERF families also existed in the rice ERF families (Supplementary Table S5 at JXB online). However, some motifs, for example CMA-5-4 in group A-5 (identical to group II), CMA-4-4, CMA-4-5, and CMA-4-8 in group A-4 (belonging to group III), CMA-2-3 in group A-2 (belonging to group IV), CMB-5-5 in group B-5 (identical to group VI), CMB-1-6, CMB-1-9, and CMB-1-10 in group B-1 (identical to group VIII), CMB-3-8, CMB-3-10, CMB-3-11, and CMB-3-12 in group B-3 (identical to IX), and CMB-4-2,CMB-4-3, and CMB-4-4 in group B-4 (identical to group X), existing in both soybean and Arabidopsis ERF families, were not found in the rice ERF family. In contrast, the motifs CMX-1 and CMX-2, identified in both the soybean and rice ERF families, were not found in the Arabidopsis ERF family.

![Fig. 1. An unrooted phylogenetic tree of soybean ERF proteins.](image-url)
Expression pattern of certain unigenes belonging to the ERF subfamily in soybean

Previous reports indicated that the roles of the A group (CBF/DREB subfamily) of transcription factors were predominantly in regulation of the abiotic stress response, while those of the B group (ERF subfamily) were involved in both biotic stress responses, abiotic stress responses, or both. Some members belonging to the B group were chosen for further study. The expression patterns of nine unigenes, namely GmERF039 in subgroup B-1, GmERF056 and GmERF057 in subgroup B-2, GmERF061, GmERF079, and GmERF069 in subgroup B-3, GmERF081 in subgroup B-4, GmERF089 in subgroup B-5, and GmERF098 in subgroup B-6, were investigated using RT-PCR under various stress conditions. As shown in Fig. 6, for high salt treatment, expression levels of GmERF056, GmERF079, and GmERF081 increased after initiation of the treatment, reached maxima at 5 h, and then decreased. The expression of GmERF039, GmERF057, and GmERF089 reached maxima at 10 h after salt treatment, and expression levels of GmERF061, GmERF069, and GmERF098 increased gradually over the 24 h period of treatment. For drought treatment, the expression levels of all nine unigenes increased gradually for at least 24 h, except for GmERF081 which peaked at 5 h. For cold treatment, the expression levels of GmERF098, GmERF081, GmERF061, and GmERF079 showed rapid increases at 5 h and remained at high levels until 24 h. The expression level of GmERF069 peaked at 5 h, and those of GmERF056, GmERF039, GmERF057, and GmERF089 were not affected. Following SMV inoculation, expression levels of GmERF039, GmERF056, and GmERF061 increased gradually for at least 24 h, whereas that of GmERF079 rapidly accumulated at 1 h, and then decreased to a low level. The expression of GmERF069 rapidly accumulated at 1 h, and was maintained for at least 10 h, but at 24 h its expression had declined to the level of uninoculated leaves. The expression level of GmERF057 reached a maximum at 2 h, and those of GmERF081, GmERF089, and GmERF098 were not affected.

Applications of ET, SA, and JA induced the expression of all nine unigenes. With ET treatment, the expression levels of all of unigenes peaked at 10 h or 24 h. For SA treatment, the expression levels were all induced and
increased gradually until at least 24 h, except that GmERF057 peaked at 2 h, GmERF079 peaked at 1 h, and GmERF069 peaked at 5 h. For JA treatment, expression of GmERF069 rapidly increased at 1 h after treatment, and remained at this high level of expression for at least 24 h. The expression of GmERF081 was rapidly accumulated until 5 h after treatment, and subsequently declined. The expression profiles of the other seven unigenes reached a high level at 10 h or 24 h after treatment. For ABA treatment, GmERF089 transcript levels were slightly increased at 2 h after treatment, and otherwise they were similar to non-stressed levels. Transcript levels of GmERF039 and GmERF081 peaked at 2 h and then returned to pre-stressed levels. Expression levels of GmERF079 and GmERF098 reached peaks at \(~\)1 h after treatment. The expression of GmERF056 peaked at \(~\)2 h, and was maintained at a high level for up to 24 h. The GmERF061 and GmERF057 transcript levels gradually increased over the entire 24 h period. GmERF069 was subjected to negative regulation by

---

**Fig. 3.** Phylogenetic relationships among the soybean ERF subfamily (group B) unigenes. Bootstrap values from 1000 replicates were used to assess the robustness of the trees. The phylogenetic tree and a schematic diagram of the protein structures of every group are shown. Each box represents the AP2/ERF domain. Conserved motifs are summarized in Supplementary Table S4 at JXB online. These motifs were defined by multiple alignments with an MEME search.
ABA, with transcription after treatment being lower than pre-treatment levels.

Overexpression of GmERF057 and GmERF089 confers increased tolerance to biotic and abiotic stresses in transgenic tobacco plants

To date, most research has focused on the B-3 subgroup of the ERF subfamily, and the functions of members from other subgroups remain largely known. To verify the functions of members of the B-2 and B-5 subgroups, GmERF057 in the B-2 subgroup and GmERF089 in the B-5 subgroup were overexpressed in tobacco plants under the control of the CaMV35S promoter. For salt tolerance analyses of GmERF057 transgenic plants, significant phenotypic differences between wild-type and transgenic plants were observed after 30 d (Fig. 7A). During that period, leaves of wild-type plants gradually lost greenness, and root elongation was severely retarded, whereas leaves of the transgenic plants remained green and root development was vigorous, indicating that transgenic plants displayed tolerance against salt stress. The effects of overexpression of GmERF057 on the response to bacterial infection are shown in Fig. 7B. The transgenic lines exhibited significantly reduced disease lesions, and bacterial numbers were significantly reduced relative to wild-type plants. Bacterial numbers in the transformed plants were ~50% of those in wild-type plants after 7 d of incubation. For GmERF089, increased drought and salt tolerances of transgenic tobacco plants were also observed in the seedling and mature growth stages, respectively (Fig. 7C, D). However, GmERF089 transgenic plants did not exhibit detectable tolerances to bacterial infection (data not shown).

Discussion

Nakano et al. (2006) systematically surveyed the gene structures, phylogeny, and conserved motifs of the ERF gene family in Arabidopsis and rice, but relatively few soybean ERF genes were studied previously. To gain further information about the ERF family in soybean, 148 members of the AP2/ERF superfamily were identified from the soybean DFCI Soybean Gene Index database, including 120 members in the ERF family. These numbers were similar to those in Arabidopsis (147 members of the AP2/ERF superfamily) and rice (157 members), and...
including 122 and 139 members, respectively, in the ERF family (Nakano et al., 2006), indicating that although soybean has a large genome of 1115 Mb (Arumuganathan and Earle, 1991) compared with Arabidopsis (145 Mb) and rice (420 Mb), the structure and phylogeny of the AP2/ERF superfamily are similar in the three species. The presence of most subgroups in the three species also suggests that many of the genes pre-date the species divergence. Likewise, some groups/subgroups are present in only one species; for example, groups XI–XIV existed only in the rice ERF family but not in the Arabidopsis or soybean ERF families (Supplementary Fig. S4 at JXB online), suggesting that these groups had evolved or been lost in one species after this divergence. However, this comparison alone provides limited functional information, whereas queries with Arabidopsis or rice ERF genes of known function could identify soybean orthologs with functional similarities. Some incompletely full-length ERF genes were missed in the present study, decreasing the likely number of ERF family members in soybean. However, a comparative analysis of soybean, Arabidopsis, and rice suggested the AP2/ERF domains were well conserved among the three species. These conserved amino acid residues probably indicate crucial roles for ERF family genes involved in different forms of physical contact with DNA. According to Allen et al. (1998), the AP2/ERF domain recognizes its target DNA via the conserved arginine and tryptophan residues in the α-sheet. Ala37 might play a crucial role in DNA binding or the stability of the AP2/ERF domain (Liu et al., 2006). The conserved motif analysis of the ERF family demonstrated that most motifs were conserved in soybean, Arabidopsis, and rice. Proteins within a subgroup that share these conserved motifs are likely to have similar functions. For example, the EAR motif is essential for gene repression (Ohta et al., 2001; Yang et al., 2005). Tobacco NtERF3, and Arabidopsis AtERF3 and AtERF4, containing the conserved EAR motif, repress the expression of a GCC-box-containing reporter gene (Fujimoto et al., 2000; Ohta et al., 2000, 2001). Mutations within the EAR motif eliminated this capacity for repression (Ohta et al., 2001). In addition to common conserved motifs in soybean, Arabidopsis, and rice, there are soybean-specific ERF

![Expression patterns of nine unigenes under various stresses. Total RNAs were isolated from soybean seedlings exposed to NaCl, drought, cold, SMV, ET, SA, Me-JA, and ABA for the indicated times. A 5 μg aliquot of total RNA was reverse-transcribed into first-strand cDNA for RT-PCR. The tubulin was amplified as a control.](image-url)
family motifs, which may have important roles in regulating biological processes in soybean; those functions need to be demonstrated further. The comparative analysis of conserved motifs in soybean, Arabidopsis, and rice suggested that protein functions have been both conserved and diverged during evolution of the ERF gene family.

Expression patterns of ERF subfamily unigenes under stress treatments

Plants undergo a range of environmental stresses in their natural environments and have evolved a wide range of mechanisms to cope with them. There are multiple stress perception and signalling pathways, some of which are

---

**Table 2. Soybean unigenes whose amino acid sequences have significant matches against the GenBank protein database**

<table>
<thead>
<tr>
<th>Generic name</th>
<th>Match description</th>
<th>GenBank accession no.</th>
<th>E-value</th>
<th>Length query (amino acids)</th>
<th>Identities</th>
</tr>
</thead>
<tbody>
<tr>
<td>GmERF005</td>
<td>DREBa transcription factor</td>
<td>AAT12423</td>
<td>7.00E-99</td>
<td>215</td>
<td>100%</td>
</tr>
<tr>
<td>GmERF006</td>
<td>DREB</td>
<td>AAP83131</td>
<td>2.00E-78</td>
<td>241</td>
<td>98%</td>
</tr>
<tr>
<td>GmERF007</td>
<td>DREB</td>
<td>AAP83131</td>
<td>2.00E-68</td>
<td>219</td>
<td>98%</td>
</tr>
<tr>
<td>GmERF023</td>
<td>DREB protein</td>
<td>AAY89658</td>
<td>4.00E-66</td>
<td>172</td>
<td>94%</td>
</tr>
<tr>
<td>GmERF025</td>
<td>Dehydration-responsive element-binding protein</td>
<td>AAP47161</td>
<td>6.00E-77</td>
<td>178</td>
<td>90%</td>
</tr>
<tr>
<td>GmERF027</td>
<td>Dehydration-responsive element-binding protein 3</td>
<td>ABB36646</td>
<td>4.00E-104</td>
<td>229</td>
<td>100%</td>
</tr>
<tr>
<td>GmERF028</td>
<td>Dehydration-responsive element-binding protein</td>
<td>AAP47161</td>
<td>3.00E-89</td>
<td>178</td>
<td>100%</td>
</tr>
<tr>
<td>GmERF029</td>
<td>Dehydration-responsive element-binding protein 3</td>
<td>AAZ0388</td>
<td>6.00E-130</td>
<td>316</td>
<td>99%</td>
</tr>
<tr>
<td>GmERF036</td>
<td>DREB2</td>
<td>AAQ57226</td>
<td>6.00E-146</td>
<td>313</td>
<td>99%</td>
</tr>
<tr>
<td>GmERF054</td>
<td>Ethylene-responsive protein</td>
<td>AAQ10777</td>
<td>0</td>
<td>383</td>
<td>97%</td>
</tr>
<tr>
<td>GmERF058</td>
<td>Ethylene-responsive protein</td>
<td>AAQ10777</td>
<td>1.00E-110</td>
<td>251</td>
<td>94%</td>
</tr>
<tr>
<td>GmERF076</td>
<td>Ethylene-responsive element-binding protein 1</td>
<td>AAM45475</td>
<td>1.00E-97</td>
<td>203</td>
<td>100%</td>
</tr>
</tbody>
</table>
specific whereas others cross-talk at various steps. This signalling cross-talk occurs in biotic stress signalling (Kunkel and Brooks, 2002), abiotic stress signalling (Chinnusamy et al., 2004), or both (Fujita et al., 2006). Recent studies have revealed ERF subfamily transcription factors as promising candidates for proteins involved in cross-talk between stress signalling pathways. In this study, the expression of nine unigenes from different subgroups of the ERF subfamily following various stress treatments was analysed. Inoculation with SMV as a biotic stress increased the transcript levels of six unigenes in soybean plants. The abiotic stresses, drought, low temperature, and high salinity, induced the expression of nine, five, and nine unigenes, respectively. The expression of all nine unigenes was induced by treatments with SA, ET, JA, and ABA. ABA is a phytohormone that is extensively involved in responses to abiotic stresses such as drought and low temperature, as well as osmotic stress. ABA also governs a variety of growth and developmental processes, including seed development, dormancy, germination, and stomatal movement. In contrast, the phytohormones SA, JA, and ET play central roles in biotic stress signalling following pathogen infection. These signalling molecules primarily regulate the protective responses of plants against both biotic and abiotic stresses via synergistic and antagonistic actions, commonly described as signalling cross-talk (Mauch-Mani and Mauch, 2005). The present results suggest that there was significant cross-talk in the expression of the nine unigenes under abiotic and biotic stress conditions. As in other studies (Park et al., 2001; Lee et al., 2004), it is speculated that cross-talk of signalling pathways in plants is a common phenomenon, allowing the formation of elaborate networks to regulate both abiotic stress tolerance and disease resistance. Hence the ERF subfamily of transcriptional factors may be connecting elements involved in cross-talk between stress signalling pathways.

Overexpression of soybean ERF subfamily genes enhanced tolerance to biotic and/or abiotic stress

The ERF subfamily genes have been characterized in tobacco (Park et al., 2001; Fischer and Droge-Laser 2004), Arabidopsis (Broun et al., 2004; Yang et al., 2005), pepper (Lee et al., 2004; Yi et al., 2004), tomato (Wang et al., 2004), corn (Chuck et al., 2002), and rice (Cao et al., 2006). Overexpression of some ERF genes enhanced resistance to biotic and abiotic stresses (He et al., 2001; Berrocal-Lobo et al., 2002; Fischer and Dröge-Laser, 2004). So far, only one ERF subfamily gene (GmEREBP1) has been isolated and characterized from soybean. The transcript abundance decreased in soybean cyst-nematode-infected roots of a susceptible cultivar, whereas it increased in infected roots of a resistant cultivar (Mazarei et al., 2002). Furthermore, ET treatment repressed GmEREBP1 mRNA accumulation in both susceptible and resistant cultivars, whereas woundings increased expression in both cultivars (Mazarei et al., 2002). GmEREBP1 transgenic soybean and Arabidopsis plants inoculated with cyst nematodes did not display significantly altered responses to nematode infection (Mazarei et al., 2007). According to the classification of Sakuma et al. (2002), GmEREBP1 belongs to the B-3 subgroup of the ERF subfamily. In the present study, GmERF057 in subgroup B-2 and GmERF089 in subgroup B-5 were further characterized by overexpression in tobacco plants. Whereas the expression of GmERF057 was induced in soybean by salinity, drought, ET, SA, JA, ABA, and SMV treatments, but not by cold stress, its expression in transgenic tobacco plants conferred enhanced tolerance to salt and pathogen stress. The expression of GmERF089 was induced by salinity, drought, ET, SA, JA, and ABA treatments, but not by cold and SMV stresses, and GmERF089 transgenic plants had enhanced tolerance to salt and drought stresses, but not to pathogen stress. The results suggested that the ERF genes in different subgroups of the ERF subfamily have distinct functions dealing with specific environmental stress conditions using both different and common signal transduction pathways. The mechanism whereby ERF subfamily genes confer different stress tolerances when subjected to biotic and abiotic stresses needs further investigation.

Supplementary data

Table S1. The primer sequences used for RT-PCR amplification of nine selected target unigenes under different stress treatments.

Table S2. Unigene list of the AP2/ERF superfamily in soybean.

Table S3. Ninety-eight ERF family unigenes in soybean.

Table S4. Summary of conserved motifs within the ERF family by comparative analysis of soybean and Arabidopsis.

Table S5. Summary of conserved motifs within the ERF family by comparative analysis of soybean and rice.

Figure S1. The deduced amino acid sequence alignment of the AP2/ERF DNA-binding domains from the 98 soybean ERF proteins in this study and 122 Arabidopsis ERF proteins described by Nakano et al. (2006) using ClustalW.

Figure S2. An unrooted phylogenetic tree of the ERF family of soybean and Arabidopsis.

Figure S3. The deduced amino acid sequence alignment of the AP2/ERF DNA-binding domains from the 98 soybean ERF proteins in this study and 139 rice ERF proteins described by Nakano et al. (2006) using ClustalW.

Figure S4. An unrooted phylogenetic tree of the ERF family of soybean and rice.
Acknowledgements

The authors are grateful to Dr RA McIntosh (Plant Breeding Institute, University of Sydney, Australia) for critical review of this manuscript. We thank Dr Lijuan Qiu (Soybean Molecular Breeding Group, Institute of Crop Sciences, CAAS) for the gift of soybean cultivars. This work was funded by the National HITECH Research and Development Program of China (‘863’ program, #2008AA10Z124 and #2006AA10A104), the National Natural Science Foundation of China (#30700508), the National Key Project for Research on Transgenic Plant and Foundation of Institute of Crop Science, Chinese Academy of Agricultural Sciences (#082060302-09).

References


