Mol Genet Genomics (2009) 281:647–664
DOI 10.1007/s00438-009-0436-8

ORIGINAL PAPER

Molecular characterization of stress-inducible *GmNAC* genes in soybean

Lam-Son Phan Tran · Truyen N. Quach · Satish K. Guttikonda · Donavan L. Aldrich · Rajesh Kumar · Anjanasree Neelakandan · Babu Valliyodan · Henry T. Nguyen

Received: 12 December 2008 / Accepted: 18 February 2009 / Published online: 11 March 2009

Abstract Drought is detrimental to plant growth and development, and often results in significant losses to the yields of economically important crops such as soybeans (*Glycine max* L.). NAC transcription factors (TFs), which consist of a large family of plant-specific TFs, have been reported to enhance drought tolerance in a number of plants. In this study, 31 unigenes that contain the complete open reading frames encoding GmNAC proteins were identified and cloned from soybean. Analysis of C-terminal regulatory domain using yeast one-hybrid system indicated that among 31 GmNAC proteins, 28 have transcriptional activation activity. Expression analysis of these *GmNAC* genes showed that they are differentially expressed in different organs, suggesting that they have diverse functions during plant growth and development. To search for the drought-inducible *GmNAC* genes, we prescreened and re-confirmed by quantitative real-time PCR analysis that nine *GmNAC* genes are induced by dehydration stress with differential induction levels in both shoot and root. The expression profiles of these nine *GmNAC* genes were also examined under other stresses such as high salinity, cold and with abscisic acid hormone treatments. Phylogenetic analysis of the GmNAC proteins with previously reported drought-inducible NAC proteins of *Arabidopsis* and rice revealed that the nine drought-inducible GmNAC proteins belong to the "stress-inducible" NAC group. The results of this systematic analysis of the GmNAC family will provide novel tools and resources for the development of improved drought tolerant transgenic soybean cultivars.

Keywords Soybean · Abiotic stress · Transcription factors · Transcriptional activity · Expression

Abbreviations

ABA	Abscisic acid
NAC	NAM, ATAF and CUC transcription factors
ORF	Open reading frame
TF	Transcription factor

Introduction

Plants grow in a dynamic and uncontrolled environment, which can frequently inflict significant constraints on growth and development. Drought stress is one of the adverse environmental factors commonly encountered by plants, and can significantly decrease crop productivity worldwide. The genetic basis of drought tolerance is not well understood, and understanding how responses to drought are regulated in planta is vital for efforts to minimize the yield impact of sub-optimal water supplies on plants.

Communicated by K. Shirasu.

Electronic supplementary material The online version of this article (doi:10.1007/s00438-009-0436-8) contains supplementary material, which is available to authorized users.

L.-S. P. Tran · T. N. Quach · S. K. Guttikonda · D. L. Aldrich · R. Kumar · A. Neelakandan · B. Valliyodan · H. T. Nguyen (✉)
Division of Plant Sciences,
National Center for Soybean Biotechnology,
University of Missouri, Columbia, MO 65211, USA
e-mail: nguyenhenry@missouri.edu
URL: http://www.expressacademic.org/~nguyenlab/

Present Address:
L.-S. P. Tran (✉)
Gene Discovery Research Group,
Signaling Pathway Research Unit, RIKEN Plant Science Center,
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan
e-mail: tran@psc.riken.jp
URL: http://labs.psc.riken.jp/spru/English/

In response to stresses such as drought, plants activate a number of defense mechanisms that function to increase tolerance to the adverse conditions imposed by these stresses (Yamaguchi-Shinozaki and Shinozaki 2006). This inducible adaptation or acclimation process has evolved throughout the plant kingdom and is critical for the survival of all plants. Identification of signaling pathways operating in stress-affected cells and the mutual interactions between these pathways is a major research effort, which has garnered a substantial amount of attention. The early events of plant adaptation to environmental stresses are the stress signal perception and subsequent signal transduction, leading to the activation of various physiological and metabolic responses (Thomashow 1999; Zhu 2002; Hazen et al. 2003; Bray 2004; Kacperska 2004; Zhang et al. 2004; Valliyodan and Nguyen 2006; Yamaguchi-Shinozaki and Shinozaki 2006; Tran et al. 2007b, c). Furthermore, a large array of genes are activated by stress conditions, meaning several proteins are produced to join the pathways which subsequently lead to synergistic enhancement of stress tolerance (Seki et al. 2003; Wang et al. 2003).

In the signal transduction networks involved in the conversion of stress signal perception to stress-responsive gene expression, various transcription factors (TFs) and *cis*-acting elements contained in stress-responsive promoters function not only as molecular switches for gene expression, but also as terminal points of signal transduction in the signaling processes. In plants, such as *Arabidopsis thaliana* and rice (*Oryza sativa*), there are more than 50 reported families of TFs (Riechmann et al. 2000; Xiong et al. 2005; Riaño-Pachón et al. 2007). Identification and molecular tailoring of novel TFs has the potential to overcome a number of important limitations involved in the generation of transgenic soybean plants with superior yield under drought conditions (Valliyodan and Nguyen 2006; Yamaguchi-Shinozaki and Shinozaki 2006; Tran et al. 2007a).

The NAM, ATAF, and CUC (NAC) TFs appeared to regulate a number of biochemical processes which protect the plants under water deficit conditions. A comprehensive analysis of the NAC TF family in *Arabidopsis* reported that there are 105 putative NAC TFs in the model plant (Riechmann et al. 2000; Ooka et al. 2003). Recently, Fang et al. (2008) systematically analyzed the NAC family in rice, and identified 140 putative NAC or NAC-like TFs. The NAC TFs are multi-functional proteins and involved in a wide range of processes such as abiotic and biotic stress responses, lateral root and plant development, flowering, secondary wall thickening, anther dehiscence, senescence and seed quality (Souer et al. 1996; Aida et al. 1997; John et al. 1997; Sablowski and Meyerowitz 1998; Xie et al. 2000; Collinge and Boller 2001; Duval et al. 2002; Hegedus et al. 2003; Vroemen et al. 2003; Fujita et al. 2004; Tran et al. 2004; Mitsuda et al. 2005; Oh et al. 2005; Uauy et al. 2006; Mitsuda et al. 2007; Yoo et al. 2007; Jensen et al. 2008).

Evidence correlating the involvement of NAC TFs with the regulation of drought stress response in plants was first reported in *Arabidopsis*. Overexpression of either *ANAC019*, *ANAC055* or *ANAC072* in *Arabidopsis* altered the expression of many stress-inducible genes in the transgenic plants, and conferred a constitutive increase in drought tolerance (Tran et al. 2004). Following this study, a number of researches have reported the potential application of NAC TFs for improvement of drought resistance in important crops. Introduction of *SNAC1* into rice enhanced the drought tolerance of these transgenic plants, as demonstrated by field studies. Transgenic rice overexpressing the *SNAC1* gene exhibited 22–34% higher seed setting than the negative control population in the field when severe drought stress was imposed during the reproductive stage (Hu et al. 2006). *OsNAC6/SNAC2* transgenic rice plants have improved drought tolerance because of enhanced expression of a large number of diverse genes encoding proteins with predicted functions related to stress tolerance (Nakashima et al. 2007). In addition to drought tolerance, these constructs demonstrated an increase in cold and salt stress tolerance as well (Hu et al. 2008).

Members of the NAC TF family, which are found exclusively in the plant kingdom, contain a highly conserved N-terminal DNA-binding NAC domain and a variable transcriptional regulation C-terminal domain (Ernst et al. 2004; Olsen et al. 2005). In *Arabidopsis,* the DNA binding site of stress-inducible NAC TFs, the NAC recognition sequence (NACRS) containing the CACG core motif, has been identified and thoroughly analyzed in planta (Simpson et al. 2003; Tran et al. 2004). The rice drought-inducible ONAC TFs bind to a similar NACRS found in rice, demonstrating that the NACRS might be conserved across the plant kingdom for stress-inducible NAC TFs (Hu et al. 2006; Fang et al. 2008). In addition to DNA binding, the NAC domain also contains a protein binding domain. The NAC domain of NAC1 TF interacts with the RING finger SINAT5, which promotes ubiquitin-related degradation of NAC1 to attenuate auxin signaling in *Arabidopsis* (Xie et al. 2002). Recently, it has also been reported that the NAC domain of ANAC TFs interacts with the zinc-finger homeodomain ZFHD1, a transcriptional activator which specifically binds to the ZFHD recognition sequence (ZFHDRS) located in the promoter region of *ERD1* gene (Tran et al. 2007a). The highly variable C-terminal domains of NAC TFs are generally considered to be the functional domains, which can act as either a transcriptional activator or repressor (Fujita et al. 2004; Tran et al. 2004; Hu et al. 2006; Kim et al. 2007). Interestingly, the C-terminal domains of relatively large NAC TFs also exhibit protein binding activity. It has been

reported that the C terminal domain of *Arabidopsis* CBNAC can interact with calmodulin proteins (Kim et al. 2007). Additionally, there is evidence which suggests that the functions of NAC TFs are dictated, to a large extent, by their interacting partners. Co-introduction of the ANAC and its interacting ZFHD1 TFs exhibited a higher degree of drought tolerance in *Arabidopsis* transgenic plants than each one individually (Tran et al. 2007a).

Soybean is a nutritionally important crop which provides an abundant source of oil and protein-rich food for worldwide human consumption. Additionally, soybean is also being viewed as an attractive crop for the production of renewable fuels such as biodiesel. Significant portions of the soybean producing areas in the U.S. have experienced severe drought over the past few years, which has resulted in significant yield reductions. So far, in soybean, only six cDNA sequences (GmNAC1-6) encoding GmNAC TFs have been identified from the soybean EST database (Meng et al. 2006). In this study, we report the identification and detailed characterization of 31 previously uncharacterized *GmNAC* genes for the purpose of selecting potential *GmNAC* gene candidates for the improvement of drought resistance in soybean via genetic engineering.

Experimental procedures

Plant growth and stress treatments

To study organ-specific expression, soybean (*Glycine max* L.) cv. Williams 82 seeds were germinated in three gallon pots containing Promix. The seedlings were grown in greenhouse conditions (28/20°C day/night temperature, photoperiod of 14/10 h, 800 μmol m^{-2} s^{-1} light intensity and 60% humidity). When the seedlings developed four nodes and three fully opened trifoliate leaves (approximately 25 days after sowing), the roots, leaves and stems were collected. Flowers and early R5 stage seeds were harvested after 55 and 100 days after sowing, respectively.

To study differential gene expression of *GmNAC* genes under different abiotic stress treatments, soybean (*Glycine max* L.) cv. Maverick seeds were germinated in pots containing a 1:1 turface to sand mixture. The seedlings were grown in greenhouse conditions with a day/night cycle of 16/8 h, 800 μmol m^{-2} s^{-1} light intensity, in 60% humidity at 28/20°C day/night temperature. For dehydration treatment, 14-day-old plants were carefully pulled out and washed to remove excess soil, then transferred onto filter paper, and allowed to dry for the indicated time periods in a growth chamber with the following controlled conditions: 60% relative humidity, 28/20°C day/night temperature and 200 μmol m^{-2} s^{-1} light intensity. Under these conditions the stem water potentials were measured to be −1.5, −8.3 and −28.6 bar for 0, 1 and 5 h dehydrated plants, respectively, using a pressure chamber (M-600, PMS Instruments, USA). For salt stress and abscisic acid (ABA) treatments, 12-day-old seedlings were carefully pulled out and washed to remove excess soil, then transferred to either distilled water (water control treatment), 250 mM NaCl or 100 μM ABA for different period of times. Low temperature treatment was conducted by transferring seedlings to water maintained at 4°C. After exposure to treatments for indicated durations, the soybean plants were immediately frozen in liquid nitrogen.

Identification and cloning of *GmNAC* genes

In order to obtain all available full-length open reading frames (ORFs) encoding GmNAC TFs, 18 representative NAC domain sequences from phylogenetically distinct subgroups of *Arabidopsis* ANAC and rice ONAC proteins (Ooka et al. 2003) were used as TBLASTN queries in Phytozome (http://www.phytozome.net), which consists of approximately 7X soybean genome coverage (estimated 98%) and 62,199 predicted transcripts (Glyma0 model). Identified ORFs were cloned using cDNA libraries prepared from total RNA extracted from 4-week-old soybean (*Glycine max* L.) cv. Williams 82 plants. cDNA libraries were synthesized using SuperScript III First-Strand Synthesis System (Invitrogen, Carlsbad, CA, USA).

Phylogenetic analysis

The amino acid sequences of NAC proteins were aligned using ClustalX (Thompson et al. 1997) with the following parameter set: gap open penalty = 10, gap extension penalty = 0.2. The alignment was adjusted manually, and unrooted phylogenetic trees were constructed by the neighbor-joining method using MEGA4 software (Tamura et al. 2007). The confidence level of monophyletic groups was estimated using a bootstrap analysis of 1,000 replicates. Only bootstrap values higher than 50% are displayed next to the branch nodes.

RNA isolation, DNase treatment and cDNA synthesis

Total RNA was isolated from plants using TRIZOL reagent (Invitrogen, Carlsbad, CA, USA). For each sample, 18 μg of total RNA was digested in a volume of 50 μL with Turbo DNA-free DNaseI (Ambion) according to the manufacturer's instructions for treatment to remove genomic DNA contamination. After DNaseI treatment, RNA concentration was determined again with the NanoDrop ND-1000 UV–Vis spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA), and first-strand cDNA synthesis was performed using 1 μg of DNAse-treated total RNA with the

iScript cDNA Synthesis Kit (Bio-Rad, Hercules, CA, USA) in a reaction volume of 20 μL according to the manufacturer supplied protocol.

Quantitative realtime-PCR

Gene-specific primers were designed using ProbeFinder Version 2.44 (https://www.roche-applied-science.com). Primer specificity was then confirmed by blasting each primer sequence against Phytozome (http://www.phytozome.net/search.php?show=blast) using BLASTN algorithm. Soybean *18S* and *CYP2* genes (Jian et al. 2008; Donavan L. Aldrich personal communication) were used as internal controls for organ-specific and stress-responsive expression studies, respectively. Quantitative RT-PCR (qRT-PCR) reactions were performed in 384-well plates (7900 HT Sequence Detection System; Applied Biosystems, Foster City, CA, USA) for all tissues and treatment conditions tested. SYBR Green PCR Master Mix (Applied Biosystems, Foster City, CA, USA) was used to prepare qRT-PCR reactions. Primer sets (0.4 μM final concentrations for each primer) were used in a final volume of 10 μL per well. The thermal profile of the qRT-PCR reactions was 50°C for 2 min, 95°C for 10 min, 40 cycles of 95°C for 15 s, and 60°C for 1 min. Dissociation curves were obtained using a thermal melting profile performed after the qRT-PCR cycles: 95°C for 15 s followed by a constant increase of the temperature between 60 and 95°C. The data obtained were analyzed with the SDS 2.2.1 software package.

Analysis of transcriptional activity in yeast

To investigate the transcriptional activity of the GmNAC proteins, the DNA fragments encoding the respective C-terminal domains were fused to the GAL4 binding domain in pGBKT7 (Clontech, Palo Alto, CA, USA) to create the fusion constructs. The AH109 yeast strain (Clontech, Palo Alto, CA, USA) was transformed with the generated constructs, and yeast transformants were streaked on selective medium lacking tryptophan, histidine and adenine to assay transcriptional activity. The vector pGBKT7 and the plasmid pGBKT7-OsNAC6 (Nakashima et al. 2007) were used as negative and positive controls, respectively.

Results

Identification, cloning and structural analysis of *GmNAC* genes

Full-length ORFs encoding putative GmNAC TFs were identified in Phytozome (http://www.phytozome.net) based on their conserved DNA binding NAC domains as described in Experimental Procedures. Initially, 31 independent putative full-length ORFs were obtained among at least 111 members of the GmNAC family identified by the Soybean Transcription Factor database (http://planttfdb.cbi.pku.edu.cn/web/index.php?sp=gm). These *GmNAC* genes were cloned using the primer sets listed in Table 1, and sequentially named GmNAC001 to GmNAC031. Out of these genes, six genes (GmNAC001 to 006) showed a high degree of similarity to the sequences AY974349 to AY974354, which had previously been cloned and deposited in GenBank (Meng et al. 2006). At the nucleotide level, *GmNAC001* and *GmNAC003* matched 100% with the GenBank sequences AY974349 and AY974351, respectively; *GmNAC002* has two differences (C_{599} to $T_{599;}$ G_{719} to A_{719}), whereas *GmNAC004* and *GmNAC005* genes have one difference (G_{847} to A_{847} for *GmNAC004* and T_{445} to G_{445} for *GmNAC005*) compared with deposited AY974350, AY974352 and AY974353 GenBank sequences, respectively. These differences led to two translational alterations in GmNAC002 (A_{200} to V_{200}; G_{240} to E_{240}) and one each in GmNAC004 (A_{243} to T_{243}) and GmNAC005 (F_{148} to V_{148}). These single nucleotide polymorphisms may be due to genotypic variation between the two cultivars used for these studies. Williams 82 was utilized for gene isolation in our research whereas Meng et al. (2006) used the Bogao genotype. As for GmNAC006, our predicted protein contains six additional amino acids flanking the N-terminal NAC domain compared with the reported AY974354 sequence. We cloned both versions, but only the longer version, with the six additional amino acids, was included in the sequence analysis. In addition, all of the isolated *GmNAC* genes showed 100% identity to the sequences predicted by Phytozome, with the exception of *GmNAC010*, *026* and *030*, which were similar, but not identical, to the predicted gene model (Table 1).

In order to examine the evolutionary relationship between the *GmNAC* genes identified in our study, a multiple alignment of their deduced proteins was performed with the ClustalX program (Supplementary Fig. 1). As expected, most of the GmNAC TFs shared a highly conserved N terminal DNA binding consistent with the typical NAC domain, which consists of five consensus subdomains and a highly variable C-terminal transcriptional regulation domain (Supplementary Fig. 1). The conserved NAC domains were then retrieved for construction of an unrooted phylogenetic tree using Cluxtal X and the neighbor-joining method in the MEGA 4 program. Phylogenetic analysis demonstrated that the majority of 31 isolated proteins belong to two main groups (I and II), which can be further classified into two subgroups. GmNAC007, 008, 020, 021, 023 and 028, postulated NAC-like proteins, form other distinct clades which may have more members yet to be identified (Supplementary Fig. 2).

Table 1 Primer pairs used in cloning of *GmNAC* genes

Name	Sequence ID[a]	Forward primer	Reverse primer
GmNAC001	Gm0129x00049	5′-GATAGATCTATGGAGAATAGAACAAGCTC-3′	5′-ATAGTCGACTTATCCTCTTTGATCATAC-3′
GmNAC002	Gm0026x00071	5′-GGCAGATCTATGGCATCAGAGCTTGAATTG-3′	5′-AACGATATCTCAAAAGGACTTGTTGGGCCAG-3′
GmNAC003	Gm0027x00010	5′-GTTAGATCTATGGGAGTTCCAGAGAGAGACCCTCTTG-3′	5′-TAAGATATCTCAGTCCCTAAACCCGAACTCAAC-3′
GmNAC004	Gm0178x00114	5′-AATAGATCTATGGGAGTTCCAGAGGAAGACCCTCTTTCC-3′	5′-TAAGATATCTCAATTCCTGAACCCGAACCCCACCG-3′
GmNAC005	Gm0083x00164	5′-CCTAGATCTATGGAAAACGTTTCGGTTC-3′	5′-GAAGATATCTCAGTAGTTCCACAAGGTAGCAAGG-3′
GmNAC006	Gm0098x00246	5′-GAGAGATCTATGGAAGACATGAACAACATGAG-3′	5′- GAGGATATCTCAGTAGAAGTTTAGAATATTGGTGGGAG-3′
GmNAC007	Gm0031x00114	5′-GGATCCATGGGGAGGTTCGAAAACCATG-3′	5′-GATATCCTAATCTAACTTGCTAGAGGAG-3′
GmNAC008	Gm0162x00055	5′-GGATCCATGGAAGTCTTAAACCCGATTG-3′	5′-GATATCTTAGGAAGCTTCACTGGTAGCAGG-3′
GmNAC009	Gm0144x00146	5′-AGATCTATGGAAAGCACCGACTCATCAAGCGG-3′	5′-GATATCCTAAGCATACCAATTCATTCCTGG-3′
GmNAC010[b]	Gm0111x00134	5′-GGATCCATGGGAACCCCACAATCCAATAAC-3′	5′-AGTACTTCACTGAAATGGAAGATATTGAGGAC-3′
GmNAC011	Gm0155x00025	5′-GGATCCATGGGAAGCCCAGAATCCAATTTG-3′	5′-AGTACTTTATCCTTGAAATTGAAGATGAGGAC-3′
GmNAC012	Gm0040x00243	5′-GGATCCATGGCCACTACAACACAACTTCAC-3′	5′-GATATCTTAATTGCAGAAGGACTTGGAGAG-3′
GmNAC013	Gm0007x00326	5′-GGATCCATGGCCGCAGCAACACAACTCCACTTACC-3′	5′-AGTACTTCAGAAGGGCCTGGAGAGGTACATGAAG-3′
GmNAC014	Gm0078x00137	5′-GGATCCATGCCAGGAGAACTCCAATTACCACC-3′	5′-GATATCTCAAAATGTCTTTGGTAGGTACGTG-3′
GmNAC015	Gm0096x00095	5′-GGATCCATGAAGGGAGAATTAGAGTTGCCAC-3′	5′-GATATCTCACATCTTCTGTAGGTACATGAACATG-3′
GmNAC016	Gm0021x00889	5′-AGATCTATGGAGAAGGTGAGTTTTGTGAAG-3′	5′-GATATCTTAAGGTTTCGTTCTAATAGTATAAG-3′
GmNAC017	Gm0161x00071	5′-GGATCCATGCCGGAAAACATGAGCATATCAGTG-3′	5′-GATATCCTACACTGACGTGTTGGACACGTGGCAC-3′
GmNAC018	Gm0180x00082	5′-GGATCCATGAGCAACATAAGCATGGTAGAGG-3′	5′-GATATCTCAGTAATTCCACACGTGGGGCATCC-3′
GmNAC019	Gm0096x00150	5′-GGATCCATGGCCAAACCAAAAATGCCAGG-3′	5′-GATATCTTACTTATCTTGGCTATCACTTCCATG-3′
GmNAC020	Gm0065x00368	5′-GGATCCATGGAAGATGGAAGCGTCAACCTTC-3′	5′-GATATCCTAAATAGGCATGCCAATTTCTTCC-3′
GmNAC021	Gm0070x00473	5′-GGATCCATGGGAGATAACAATGTCAACCTTCCACC-3′	5′-GATATCCTAATAATTTGGTAAGCTAATTTCATC-3′
GmNAC022	Gm0129x00097	5′-GGATCCATGATGGAAGAGAGAAATAATGATGC-3′	5′-GATATCCTAAGTATAACACTTAGTTGTGG-3′
GmNAC023	Gm0118x00149	5′-GGATCCATGGAAGATCATCCACCTGGGTTTCG-3′	5′-GATATCTTAGAGCCAATTAAATTGTTCCC-3′
GmNAC024	Gm0087x00031	5′-AGATCTATGGGTGAGGCTTCGGGAGCTGG-3′	5′-GATATCTCAAGGAGAAATACTTCTACCTACTAG-3′
GmNAC025	Gm0065x00230	5′-AGATCTATGCCCGGTTTGTCGGTGGTACAGAC-3′	5′-AGTACTTTAAAATTTAAGATATCCCCAAATGTC-3′
GmNAC026[b]	Gm0096x00303	5′-AGATCTATGGGTTCGGTGGACTGTTATCCTTC-3′	5′-GATATCTTAGATTGTTAGATATGACACCGGGAG-3′
GmNAC027	Gm0015x00210	5′-TGATCAATGGGTGCCGTCGTCGAGTGTTACCC-3′	5′-GATATCTTAAGATCTGACATATGCCCATATATTAAC-3′
GmNAC028	Gm0070x00216	5′-GGATCCATGACATGGTGCAATGACTCTGATGAGG-3′	5′-GATATCCTAGAATCTCTTCCCAGTGTTACTTTC-3′
GmNAC029	Gm0054x00192	5′-GGATCCATGAGCAACATAAGCATGGTAGAGG-3′	5′-GATATCTCAGTAATTGTTCCACATGTGGGGC-3′
GmNAC030[b]	Gm0202x00002	5′-GGATCCATGCCGGAAAACATGAGCATATCAGTG-3′	5′-GATATCCTACACTGACGTGTTGGACACGTGGCAC-3′
GmNAC031	Gm0065x00353	5′-GGATCCATGGCCAAACCAAAAATGCCAGG-3′	5′-ATGATATCTTACTTATCTTGGCTACCACTTCC-3′

[a] Sequence ID as based on Glyma0 model (http://www.soybeangenome.org/documents/Glyma0.1.cds.fa.txt)

[b] Similar to gene model predicted by Glyma0 model

◀ **Fig. 1** Analysis of transcriptional activity of 31 GmNAC TFs. The cDNA fragments encoding C-terminal regions of 31 GmNAC TFs (TRR) were fused with GAL4 DNA binding domain in the pGBKT7 yeast vector. These generated constructs were transformed into the AH109 yeast reporter strain, and the ability of each transformant to grow on medium lacking histidine and adenine was assessed to analyze the transcriptional activity of the fusion proteins

Transcriptional activity of GmNAC TFs

The NAC TFs are classified by the conserved N-terminal DNA-binding domain and the variable C-terminal transcriptional regulation region (TRR), which can act as either a repressor or activator in regulatory pathways (Tran et al. 2004; Kim et al. 2007). Thus, to further characterize their function, we investigated the transcriptional activity of GmNAC TFs by fusing cDNA fragments containing the C-terminal domains of the 31 GmNAC TFs with the GAL4 DNA-binding domain of yeast expression vector pGBKT7 for transcriptional activation assay using a yeast one-hybrid system. Detailed information on the primer sets used in this study is provided in Table 2. The generated constructs, positive control (plasmid pGBKT7-OsNAC6) and negative control (vector pGBKT7) were individually transformed into the AH109 yeast strain to assay their ability to activate transcription of the *HIS3* and *ADE2* reporter genes from the GAL4 upstream activation sequence. As shown in Fig. 1, all yeast transformants, including those containing only the vector control pGBKT7, could grow normally on the selection medium lacking tryptophan (-Trp) which was used for the selection of transformants (normal growth control). On the other hand, when selective medium lacking histidine and adenine was used to assess the transcriptional activity (-Trp/-His/-Ade), yeast cell harboring the C-terminal domains of GmNAC007, 008, 016, 028 and vector pGBKT7 could not grow, whereas the transformants carrying C-terminal domains of the remaining 27 GmNAC TFs and the OsNAC6 were able to grow well, thereby indicating that, with the exception of the GmNAC007, 008, 016 and 028 TFs, the rest of the GmNAC TFs investigated in this study possess the potential activation domain in the yeast assay system, i.e. they function as transcriptional activators (Fig. 1).

Expression of *GmNAC* genes in various plant organs

A number of reports have suggested that organ-specific NAC TFs play an important role in plant growth and development (Xie et al. 2000; He et al. 2005; Mitsuda et al. 2007; Yoo et al. 2007). Root development is one of the key traits correlated to mechanisms of drought resistance (Sharp et al. 2004). Organ-specific gene expression profiling enables the determination of differentially expressed

Table 2 Primer pairs used in analysis of transcriptional activity

Name	Forward primer	Reverse primer	Size of AD (aa)
GmNAC001-TRR	5′-GATGAATTCAAGAATATTGGAAAATCAATGG-3′	5′-ATAGATCTTTATCCTCTTTGATCATAC-3′	118
GmNAC002-TRR	5′-GATGAATTCAAGGGCACGATCGAGAAACTG-3′	5′-ATAGATCTTCAAAAGGACTTGTTGGGCCAG-3′	142
GmNAC003-TRR	5′-GATGAATTCAACTCAAGTTCACAAAAGGTTGAG-3′	5′-ATAGATCTTCAGTCCCTAAACCCGAACTCAAC-3′	173
GmNAC004-TRR	5′-AATAGATCTGTAACTCGAGTGCACAGAAGGC-3′	5′-ATGTCGACTCAATTCCTGAACCCGAACC-3′	183
GmNAC005-TRR	5′-GATGAATTCAGTTCAGGTGTGAAAAGGACTC-3′	5′-ATAGATCTTCAGTAGTTCCACAAGGTAGC-3′	189
GmNAC006-TRR	5′-GATGAATTCGCAACTTCGTTGAAAGTGC-3′	5′-ATAGATCTTCAGTAGAAGTTTAGAATATTGG-3′	136
GmNAC007-TRR	5′-ATGGATCCGTAACAAGATGAAGACAAG-3′	5′-ATCTGCAGCTAATCTAACTTGCTAG-3′	64
GmNAC008-TRR	5′-ATGGATCCGTAAGACGAAGAAGGAAGG-3′	5′-ATCTGCAGTTAGGAAGCTTCACTGG-3′	58
GmNAC009-TRR	5′-ATAGATCTGTAGCAACACGCATAGGTC-3′	5′-ATCTGCAGCTAAGCATACCAATTC-3′	198
GmNAC010-TRR	5′-ATGGATCCGTTTCAAGCATGCTGTATC-3′	5′-ATCTGCAGTCACTGAAATGGAAGATATTG-3′	197
GmNAC011-TRR	5′-ATGGATCCGTTCCAAACATGCTCTAAC-3′	5′-ATCTGCAGTTATCCTTGAAATTGAAG-3′	188
GmNAC012-TRR	5′-ATGGATCCGTAAAGGTGCAATTGAAAAGC-3′	5′-ATGTCGACTTAATTGCAGAAGGACTTG-3′	121
GmNAC013-TRR	5′-ATGGATCCGTAAAGGTGCAATTGAGAAGC-3′	5′-ATCTGCAGTCAGAAGGGCCTGGAGAG-3′	109
GmNAC014-TRR	5′-ATGGATCCGTAAAGGGAAGATTGAG-3′	5′-ATCTGCAGTCAAAATGTCTTTGGTAG-3′	137
GmNAC015-TRR	5′-ATGGATCCGTAAAGGGAAGATTGAGAAATAC-3′	5′-ATCTGCAGTCACATCTTCTGTAGGTACATG-3′	136
GmNAC016-TRR	5′-ATGGATCCGTAAGAGAAGCGGTGCTAAAAATG-3′	5′-ATCTGCAGTTAAGGTTTCGTTCTAATAG-3′	68
GmNAC017-TRR	5′-ATGGATCCGTAAGAACCATCTGAAAACC-3′	5′-ATCTGCAGCTACACTGACGTGTTGGA-3′	268
GmNAC018-TRR	5′-ATGGATCCGTAACAGTGAAGTTTTAGC-3′	5′-ATCTGCAGTCAGTAATTCCACACGTGG-3′	144
GmNAC019-TRR	5′-ATGGATCCGTGAAGGTCCTGGTCCTAG-3′	5′-ATCTGCAGTTACTTATCTTGGCTATC-3′	348
GmNAC020-TRR	5′-ATGGATCCGTTGTTCCTCTGGGTTTG-3′	5′-ATCTGCAGCTAAATAGGCATGCC-3′	75
GmNAC021-TRR	5′-ATGGATCCGTTCAGATTCTTCTTCCTC-3′	5′-ATCTGCAGCTAATAATTTGGTAAGC-3′	63
GmNAC022-TRR	5′-ATGGATCCGTATCAATGCTACAGCTC-3′	5′-ATCTGCAGCTAAGTATAACACTTAG-3′	265
GmNAC023-TRR	5′-ATGGATCCGTGTCATATCGGGGAGCTTTC-3′	5′-ATCTGCAGTTAGAGCCAATTAAATTG-3′	88
GmNAC024-TRR	5′-ATGGATCCGTAGTGGACCTGGTCC-3′	5′-ATCTGCAGTCAAGGAGAAATACTTC-3′	380
GmNAC025-TRR	5′-ATAGATCTGTCATGATGAGAGCCTTG-3′	5′-ATGTCGACTTAAAATTTAAGATATCC-3′	431
GmNAC026-TRR	5′-ATAGATCTGTCATGATGAGAGCCTTGAAG-3′	5′-ATCTGCAGTTAGATTGTTAGATATGAC-3′	201
GmNAC027-TRR	5′-ATTGATCAGTCAAGATGAGAGTCTTG-3′	5′-ATGTCGACTTAAGATCTGACATATGC-3′	418
GmNAC028-TRR	5′-ATGGATCCGTTATCAGACACAACCTAG-3′	5′-ATCTGCAGCTAGAATCTCTTCCCAGTG-3′	102
GmNAC029-TRR	5′-ATGGATCCGTAACAGAGAAGTTTCAGCC-3′	5′-ATCTGCAGTCAGTAATTGTTCCACATGTGGGGC-3′	145
GmNAC030-TRR	5′-ATGGATCCGTAAGAACCATCTGAAAACC-3′	5′-ATCTGCAGCTACACTGACGTGTTGG-3′	225
GmNAC031-TRR	5′-ATGGATCCGTGAGGGTCCTGGTCCTAGG-3′	5′-ATCTGCAGTTACTTATCTTGGCTACC-3′	328

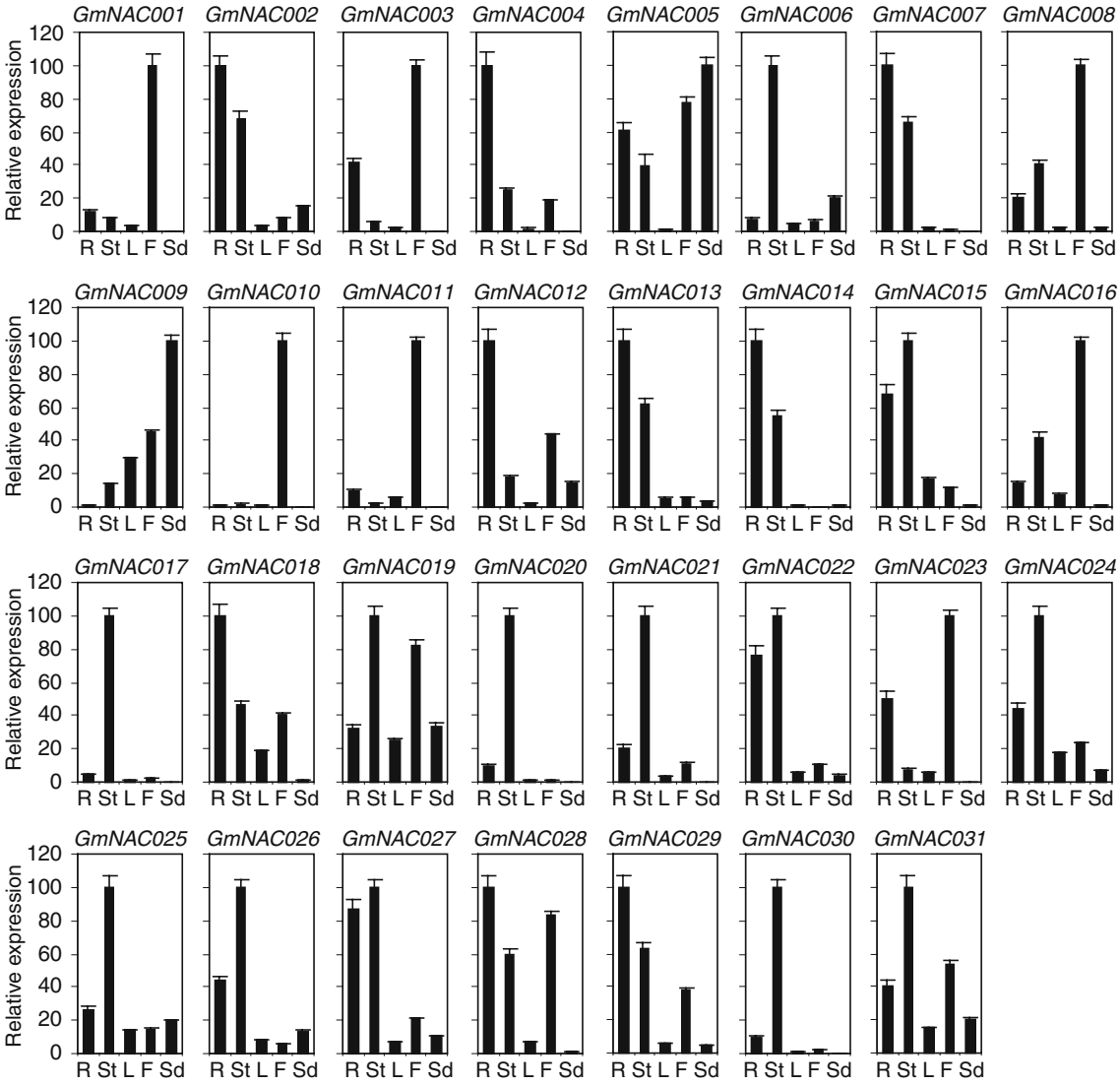

Fig. 2 Expression of 31 *GmNAC* genes in different tissues. The mean expression of 31 *GmNAC* genes was examined in five types of tissues: root (*R*); stem (*St*); leaves (*L*); flowers (*F*) and seeds (*Sd*). The highest expression level was normalized to a value of 100. The standard errors of three technical replications are shown

genes and provides clues about candidate genes that can potentially be exploited to enhance drought resistance when overexpressed in transgenic plant systems. To accomplish this, expression profiling utilizing qRT-PCR was performed using the gene-specific primers sets provided in Table 3 to determine the expression of the 31 *GmNAC* genes in five different organs of soybean, namely the roots, stems, leaves, flowers and immature seeds (R5 stage). Our results demonstrate that many of the isolated 31 *GmNAC* genes exhibit a propensity toward organ-specific expression. Thirteen *GmNAC* genes, *GmNAC006*, *015*, *017*, *019-022*, *024-027*, *030* and *031* showed high expression in stems relative to all other organs. Nine *GmNAC* genes, *GmNAC002*, *004*, *007*, *012*, *013*, *014*, *018*, *028* and *029* were expressed primarily in roots, while 7 *GmNAC* genes, *GmNAC001*, *003*, *008*, *010*, *011*, *016* and *023* appeared to be flower-specific. Transcripts of the remaining *GmNAC5* and *GmNAC9* were mainly detected in seeds (Fig. 2). The different expression profiles of these *GmNAC* genes suggest that their products are involved in functionally diverse roles during soybean growth and development.

Screening for dehydration-inducible *GmNAC* genes

Currently, the biological functions of these isolated GmNAC TFs remain unknown. Our major goal was to identify the dehydration-inducible GmNAC TFs that can potentially be used to engineer soybean plants with improved drought resistance. In order to find the dehydration-inducible genes among our 31 isolated *GmNAC* genes, we

Table 3 Primer pairs used in quantitative RT-PCR

Name	Forward primer	Reverse primer
GmNAC001	5′-TCCATGAGGCTAGATGACTGG-3′	5′-TGCCTCCATTGATTTTCCA-3′
GmNAC002	5′-TCGGAGCAGGTGGTATCG-3′	5′-GGCTTTTCTCCCACTCGTT-3′
GmNAC003	5′-ACCCTGCCACGAGTGAAC-3′	5′-GTTCGGAGACCCGAATTTCT-3′
GmNAC004	5′-ACGTCAGTTCCGCAAAAGAT-3′	5′-GGACCCGTTGGTTTCTCAC-3′
GmNAC005	5′-TGGGTAAAAGGCACTTTGAGA-3′	5′-AGGTCCATTGGTGCAACTG-3′
GmNAC006	5′-TACAAATGCCCTTTGGGAAG-3′	5′-TTGGGTCCAATCCGAGAG-3′
GmNAC007	5′-TGCGCCTTGTAGCTTCCTA-3′	5′-TGAAATATGCGATACACAGCAA-3′
GmNAC008	5′-AAGAAAGAACGTTTTGGCTTTCT-3′	5′-AATTCATGCATCACCCAATTC-3′
GmNAC009	5′-TGTTGTGAGGGATGATCAAGA-3′	5′-AGTTGATTAAGCAGAGTGGCAAT-3′
GmNAC010	5′-CCCTTTTACCAATTTTGAAGAATAAC-3′	5′-GATTTCTGAGACATGAGTATTGTGG-3′
GmNAC011	5′-AGGGGTTGGATCAAGTGCT-3′	5′-TGATTGGAGAGCATTGCTGA-3′
GmNAC012	5′-AAAGGTGCAATTGAAAAGCAA-3′	5′-CCTCCATTTCGGAACATTCA-3′
GmNAC013	5′-GCCACTCTCGGGAATAATCA-3′	5′-GCCTGGAGAGGTACATGAAGA-3′
GmNAC014	5′-CCAAATCAACCATCTCTCACC-3′	5′-CAAAATGTCTTTGGTAGGTACGTG-3′
GmNAC015	5′-ATATCGCCTCGCCAATGTT-3′	5′-CCCAATCATCAAGCCTCAAG-3′
GmNAC016	5′-TGAAGAAGAGAAGCGGTGCTA-3′	5′-CCTTAGAGTTCCTCACCTTGTTG-3′
GmNAC017	5′-AGGGGTTGCAAGGAGGAG-3′	5′-TGCAAACCTTCCATAGCTGTT-3′
GmNAC018	5′-AAGTATGCAACAGGGTTACGC-3′	5′-CTGTGGCCTTCCAATACCC-3′
GmNAC019	5′-CCAGAATTGGTGCCCTCTT-3′	5′-TGCCAACGTCAGCTTGATTA-3′
GmNAC020	5′-GCGATGATGATGACAGTGGA-3′	5′-TCATCATCCAACGACAGATAAACT-3′
GmNAC021	5′-TGGGTGATATGTCGTGTTTATGA-3′	5′-AGCACGACAGCTCTGTTCC-3′
GmNAC022	5′-CCTCAGAAGGTTCCAAGTGC-3′	5′-TCCTTTGGCAGCTCTACCTT-3′
GmNAC023	5′-ACCTGGGTTTCGTTTCTTCC-3′	5′-TGGCCTTCTAGCTTGTTGTG-3′
GmNAC024	5′-CAACATTGCTGCATCAGGTAA-3′	5′-AGTCCCTTTCTGCCAGCTCT-3′
GmNAC025	5′-GAGTCAACCGGCAAAAGAAA-3′	5′-GCACCCTTCTGAGCAAGAAA-3′
GmNAC026	5′-GCTTCCACTCCTGTGGCTTA-3′	5′-TGCAACAACAGCCAGATCA-3′
GmNAC027	5′-TTCCACTGTTACAGATGATGTGG-3′	5′-TTTCTTAGGTCGGTTGACTCG-3′
GmNAC028	5′-CTGCAACTCTTTTGGATTGCT-3′	5′-GACTCACATGGACGTGATCATAA-3′
GmNAC029	5′-GCTGCTATGAGGACACAGGAT-3′	5′-TGGTCAAAACTTATGTAAGAGTCCA-3′
GmNAC030	5′-CCATCTGAAAACCCTAGACAGC-3′	5′-TCGAACAAGTGGCTCCTTCTA-3′
GmNAC031	5′-GAATTGATGCCCTCATTCTCA-3′	5′-TACAGAGCCAACTTGATTGCTT-3′
18S	5′-CCTGCGGCTTAATTTGACTCAAC-3′	5′-TAAGAACGGCCATGCACCA-3′
CYP2	5′-CGGGACCAGTGTGCTTCTTCA-3′	5′-CCCCTCCACTACAAAGGCTCG-3′

employed qRT-PCR to determine the expression levels of all 31 *GmNAC* genes in 14-day-old soybean plants subjected to 10 h dehydration stress. The specific primer sets used for qRT-PCR analysis can be obtained on Table 3. We observed that out of 31 genes, nine *GmNAC* genes, namely *GmNAC002, 003, 004, 010, 012, 013, 015, 020* and *028* displayed apparent induction under dehydration (data not shown). Expression of these nine *GmNAC* genes was then analyzed in a more rigorous fashion using 14-day-old soybean plants subjected to varying durations of dehydration treatment. Dehydration-inducible expression of these *GmNAC* genes was clearly confirmed as shown in Fig. 3. The transcript accumulation of the *GmNAC* genes started after 1 h of dehydration stress when stem water potential of the plants was −8.3 bar. The water potentials of the control plants and 5 h dehydrated plants were measured to be −1.5 and −28.6 bar, respectively, indicating an equivalency of 5 h dehydration to severed water stress. Additionally, inducibility of these 9 *GmNAC* genes by dehydration was further confirmed in tissues collected from 12-day-old soybean plants which had been dehydrated for various time periods (data not shown). All nine dehydration-induced *GmNAC* genes, especially *GmNAC003* and *GmNAC004*, whose expression is most strongly induced by dehydration, are considered to be good candidates for transgenic studies in soybean.

Since many drought resistance mechanisms are associated with traits related to shoot and root, thus as a the first

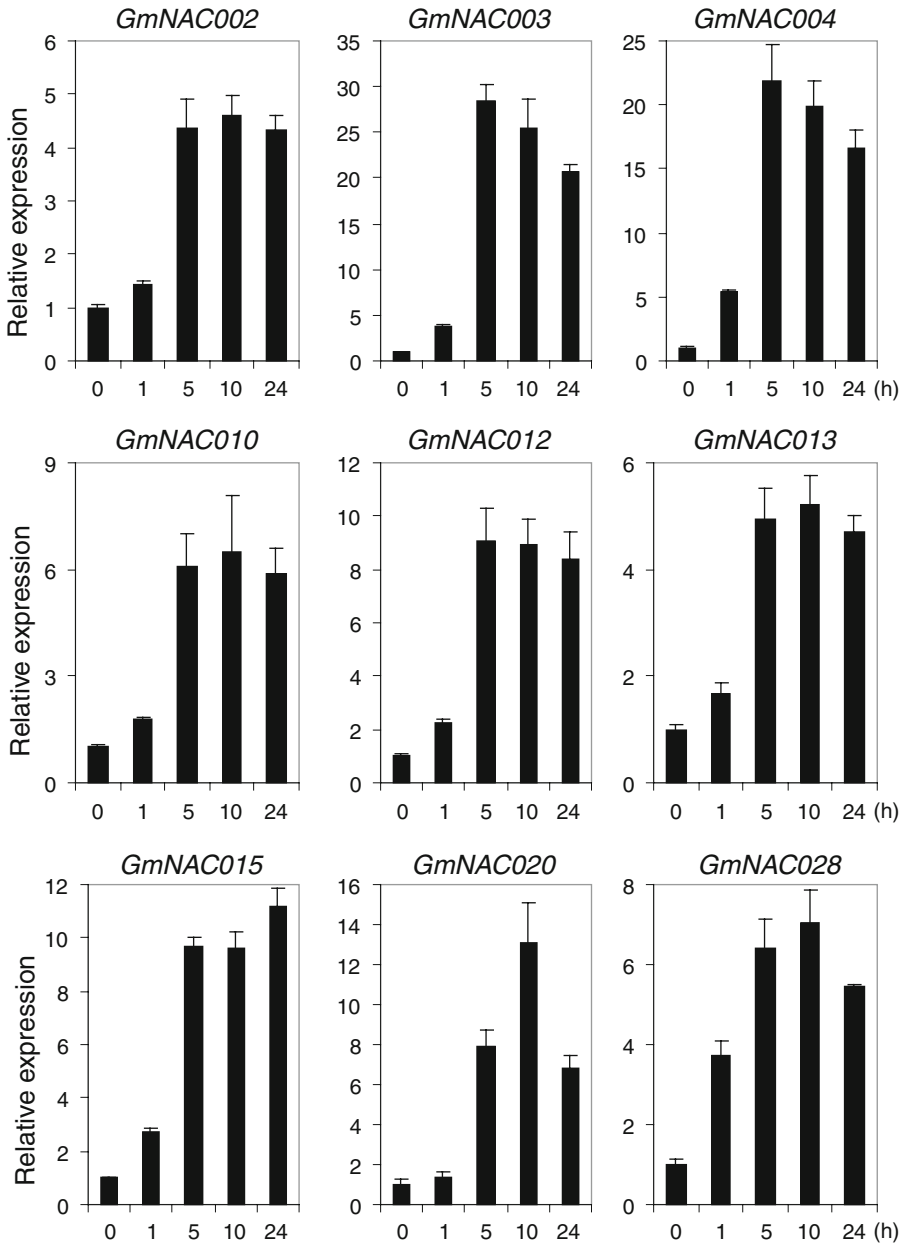

Fig. 3 Identification of nine dehydration-inducible *GmNAC* genes. Nine *GmNAC* genes were identified after expression levels of 31 *GmNAC* genes were prescreened in 0 and 10 h dehydrated 14-day-old soybean plants by qRT-PCR. Stress-induced expression of these 9 *GmNAC* genes was confirmed by qRT-PCR in 14 day soybean plants which had been dehydrated for indicated durations. Mean relative expression levels were normalized to a value of 1 in unstressed plant (0 h). The standard errors of three biological replications are shown

step to understand the mode of action of drought-inducible GmNAC TFs, the expression patterns of the identified stress-inducible *GmNAC* genes were screened for organ-specific stress induction using isolated roots and shoots tissues collected from 12-day-old soybean plants subjected to dehydration stress for various time periods. As shown in Fig. 4 the transcript levels of *GmNAC003*, *GmNAC015* and *GmNAC020* genes increased significantly in roots, whereas those of *GmNAC002*, *GmNAC010*, *GmNAC013*, *GmNAC028* increased in shoots. While *GmNAC004* and *GmNAC012* transcripts accumulated markedly in both root and shoot tissues.

Inducibility of dehydration-inducible *GmNAC* genes by other stress and hormone treatments

Molecular and biochemical studies have suggested a model that consists of both ABA-dependent and ABA-independent pathways for regulating gene expression in response to dehydration stress (Yamaguchi-Shinozaki and Shinozaki 2005, 2006). Additionally, regulatory pathways modulating responses to different stresses are usually overlapped, especially those implicated in dehydration and salt stress response (Seki et al. 2003; Yamaguchi-Shinozaki and Shinozaki 2005). Therefore, determining whether these nine dehydration-

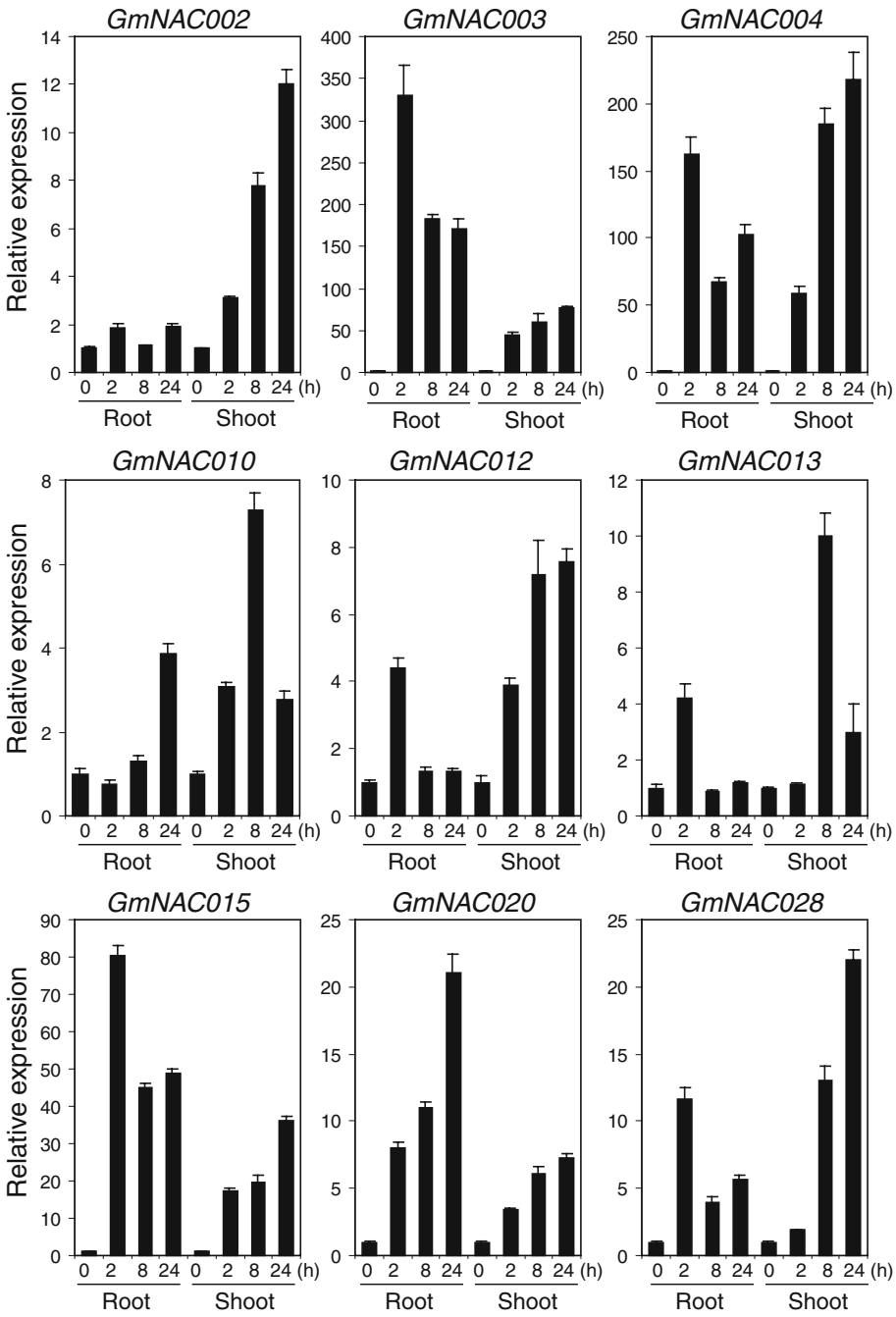

Fig. 4 Tissue related inducibility of nine dehydration-inducible *GmNAC* genes. Tissue related inducibility of dehydration-inducible *GmNACs* was examined by qRT-PCR in shoot and root tissues harvested from 12-day-old soybean plants, dehydrated for indicated durations. Mean relative expression levels were normalized to a value of 1 in unstressed plant (0 h). The standard errors of three technical replications are shown

inducible GmNAC TFs behave in an ABA-dependent or independent manner can provide further insight into their functions in planta. Thus, we investigated the expression of dehydration-inducible *GmNAC* genes under various stress and ABA treatments. Figure 5 shows that the dehydration-inducible *GmNAC* genes were also induced by high salinity, with the exception of *GmNAC020*. In addition, transcript levels of two genes, *GmNAC003* and *GmNAC004*, were also increased under exogenous ABA treatment, indicating that these two genes may function in drought stress through an ABA-dependent pathway. On the other hand, *GmNAC002*, *003*, *004*, *013*, *015* and *028* were also induced by cold stress. Hence, the nine stress-responsive *GmNAC* genes exhibit a large diversity in stress-induced expression, suggesting that they may participate in the regulation of a wide spectrum of responses to different abiotic stresses.

Comparison and phylogenetic analysis of the stress-inducible GmNAC, ANAC and ONAC proteins

To further predict the function of the dehydration induced GmNAC proteins, a sequence comparison of the dehydration-inducible GmNAC, ANAC (*Arabidopsis* NAC) and ONAC

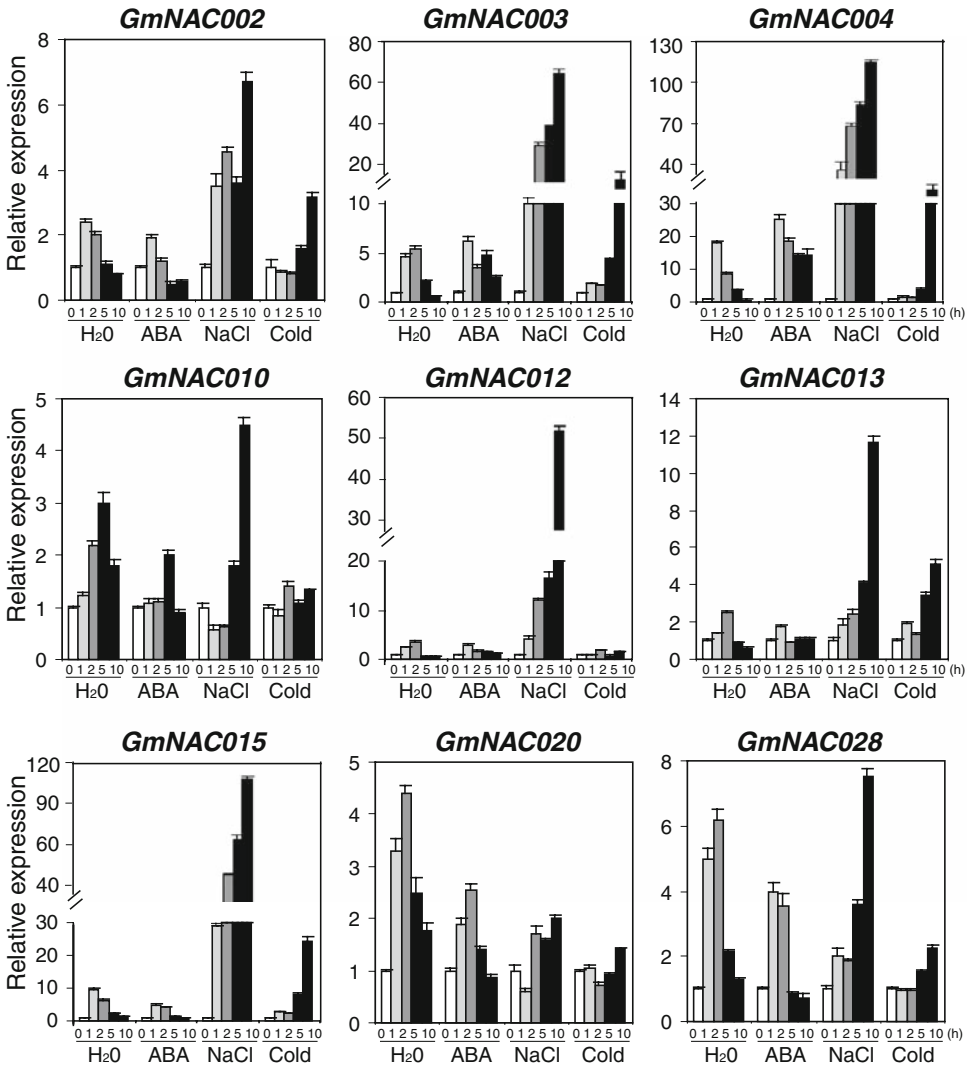

Fig. 5 Inducibility of dehydration-inducible *GmNAC* genes by other stress and ABA treatments. Expression of dehydration-inducible *GmNAC* genes was studied by qRT-PCR in 12 days soybean plants which had been transferred for hydroponic growth at 28°C in water (H_2O) for control, 100 mM ABA (ABA), 250 mM NaCl (NaCl) or at 4°C (Cold) for indicated durations. Mean relative expression levels were normalized to a value of 1 in unstressed plant (0 h). The standard errors of three technical replications are shown

(*Oryza* NAC) proteins was performed using Cluxtal X. With the exception of GmNAC020 and 028, all of these stress-responsive proteins share substantial sequence similarity in the N-terminal domains with five typical motifs. These proteins also contain a putative nuclear localization signal, suggesting that they localize in the nucleus. Interestingly, we found that ANAC002, GmNAC002, GmNAC012 and GmNAC013 share a high sequence similarity, not only in the N-terminal NAC domains, but also in their C-terminal domains (Fig. 6). The stress-responsive ANAC and ONAC proteins were then included into a phylogenetic analysis of the GmNAC proteins on the basis of similarities in their NAC domains to examine the evolutionary relationship of the stress-inducible NAC proteins among different species. As shown in Fig. 7, all stress-responsive ANACs, GmNACs and ONACs grouped into a stress-responsive subgroup, except for GmNAC020 and GmNAC028, which are located into two distinct clades.

Discussion

Increasing evidence suggests that NAC TFs play an important role in the enhancement of drought tolerance for a number of plants (Tran et al. 2004; Hu et al. 2006, 2008). To search for *GmNAC* genes for the purpose of improvement of drought resistance in soybean by genetic engineering, we have performed a homology search against the publicly available soybean sequence database (http://www.phytozome.net) using 18 representative DNA binding NAC domain sequences (Ooka et al. 2003). Structural analysis of the cloned GmNAC proteins indicate that most of them possess the highly conserved N-terminal DNA binding NAC domain, which contains five typical motifs (Supplementary Fig. 1) (Ooka et al. 2003). This type of NAC domain appears to be representative of the majority of NAC proteins across the plant kingdom. Recently, Fang et al. (2008) systematically analyzed NAC members from rice,

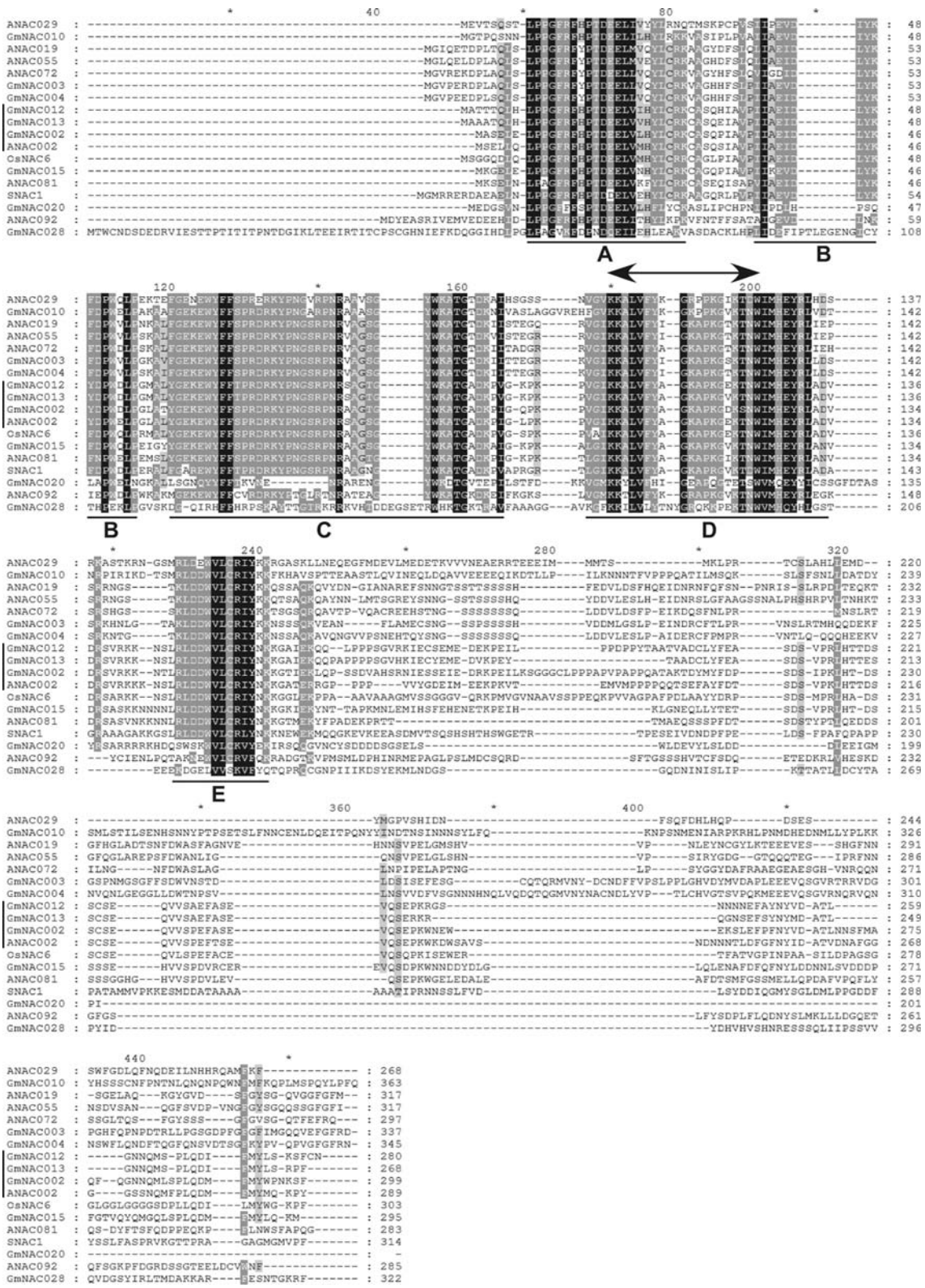

Fig. 6 Comparison of the amino acid sequences of stress-inducible GmNAC, ANAC and ONAC proteins. Identical amino acids are indicated by *white letters* on a *black background*. The consensus NAC subdomains (A–E) are indicated by *underlines*. The putative nuclear localization signal is shown by a *double-headed arrow* above the sequence. Highly similar NAC proteins are indicated by a *vertical sideline*

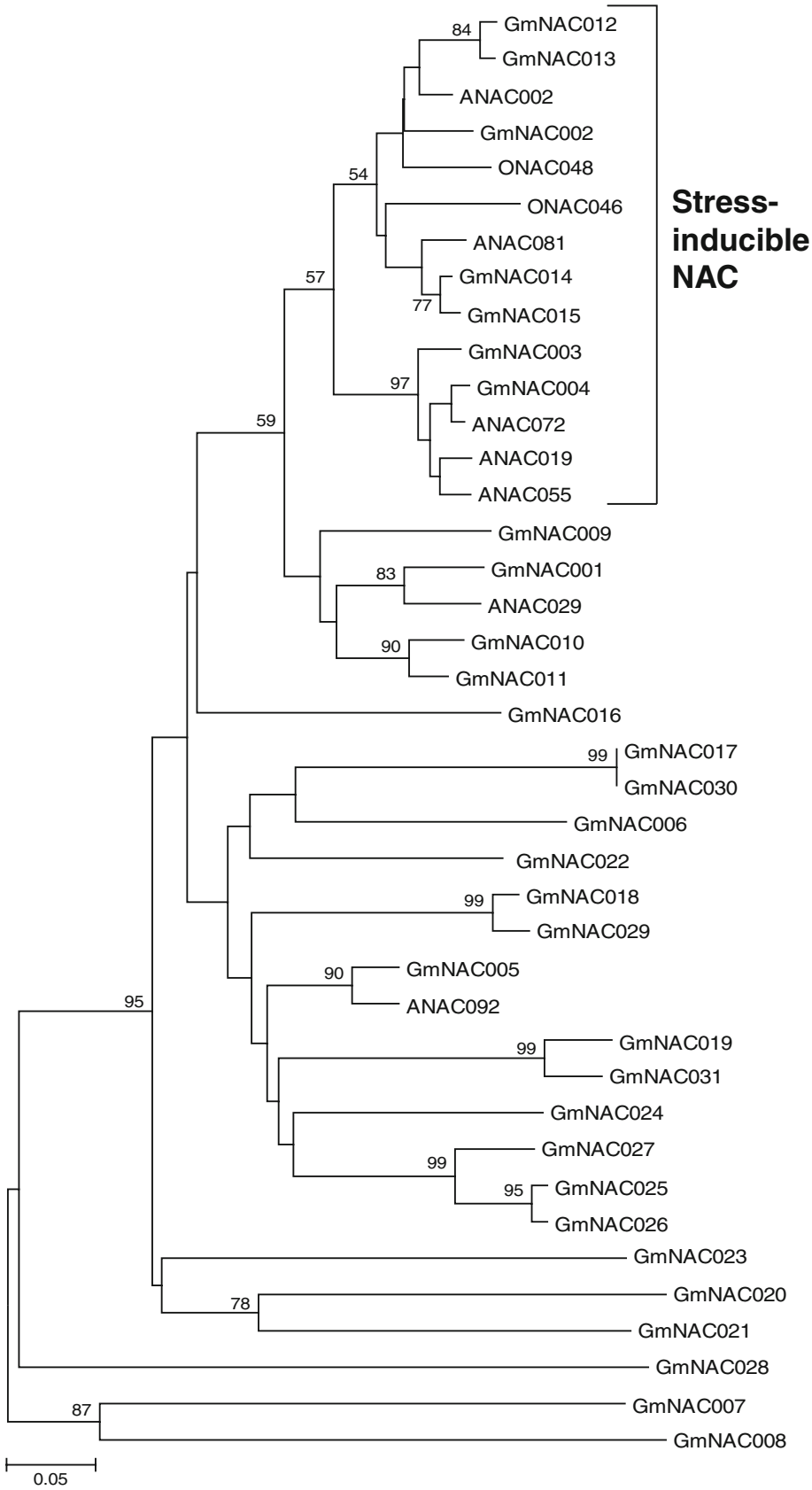

Fig. 7 Evolutionary comparison of GmNACs with the stress responsive ANAC and ONAC proteins. The conserved N-terminal DNA binding NAC domains were used to construct the unrooted phylogenetic tree. The *bar* indicates the relative divergence of the sequences examined. Bootstrap values higher than 50% are displayed next to the branch nodes

and reported that out of 140 NAC and NAC-like TFs, 97 ONAC TFs shared the typical NAC domain containing the five consensus motifs within. They suggested that those members whose DNA-binding NAC domains are not comprised of these five typical motifs should be more properly referred to as "NAC-like" proteins. Thus, among the 31 GmNAC TFs, GmNAC007, 008, 020, 021, 023, 025 and 028 are perhaps more accurately classified as NAC-like proteins (Supplementary Fig. 1). Furthermore, several GmNAC proteins, GmNAC003 and 004, GmNAC014 and 015, GmNAC025 and 026, and GmNAC017 and 030 appear to be isoforms, as they exhibit considerable homology to one another, perhaps due to genome duplication (Supplementary Fig. 1) (Shoemaker et al. 2006). To further characterize the GmNAC TFs, their transcriptional activity, which is conferred by the C-terminal TRR, was investigated using a yeast system. The results indicate that out of the 31 GmNAC TFs, 27 function as transcriptional activators (Fig. 1). Since the GmNAC007, 008, 016 and 028 displayed no transcriptional activity in yeast, they may function as transcriptional repressors (Fig. 1). Alternatively, the TRR-s of these four GmNAC proteins may contain negative regulatory domains as reported for AREB1 and DREB2A TFs. Removing the negative regulatory domain from either AREB1 or DREB2A significantly increased their transcriptional activation activity (Fujita et al. 2005; Sakuma et al. 2006).

Organ-specific expression analysis of the *GmNAC* genes indicates that the functions of the isolated GmNAC TFs are very diversified. Most of the *GmNAC* genes examined showed significant expression in only one to two organs (Fig. 2). The organ-specifically expressed genes deserve further investigation regarding their biological function. *NAC1* of *Arabidopsis* has been shown to be expressed mainly in roots, and functional analysis has proven that NAC1 plays a major role in the regulation of lateral root development (Xie et al. 2000). Expression of *AtNAC2* gene was observed at its highest level in root tissues, and overexpression of *AtNAC2* in *Arabidopsis* stimulated the development of lateral roots in the transgenic plants (He et al. 2005). In our group, special attention has been given to nine root-related *GmNAC* genes, namely *GmNAC002*, *004*, *007*, *012*, *013*, *014*, *018*, *028,* and *029* as one of our objectives is to study the regulation of root growth and development under water stress conditions.

A great deal of evidence has demonstrated that overexpression of dehydration-inducible NAC TFs has the potential to enhance drought resistance in plants (Tran et al. 2004; Hu et al. 2006, 2008; Nakashima et al. 2007). Expression of nine *GmNAC* genes, namely *GmNAC002, 003, 004, 010, 012, 013, 015, 020,* and *028* was found to be induced under dehydration (Fig. 3), suggesting that they are potential candidates for improvement of drought resistance in soybean, and their respective promoters may also be used for inducible expression of engineered transgenes in genetically modified plants for conferring stress tolerance. Constitutive overexpression of stress-inducible genes is often associated with growth retardation, with the degree of retardation being correlated with the expression level of the transgenes (Kasuga et al. 1999; Aharoni et al. 2004). To overcome this problem, stress-inducible promoters are typically used for the genetic engineering of enhanced stress resistance of important crops (Kasuga et al. 1999; Bhatnagar-Mathur et al. 2007; James et al. 2008).

Mechanisms of drought resistance are usually associated with major physiological traits, and occur primarily in shoot and root (Sharp et al. 2004; Hu et al. 2006). Expression of *SNAC1* is induced specifically in guard cells under dehydration stress, and consequently, overexpression of this gene in rice promoted stomata closure, thereby reducing water loss and led to improved drought resistance of the transgenic plants (Hu et al. 2006). On the other hand, overexpression of the *Arabidopsis HDG11* gene in transgenic tobacco, which encodes a homeodomain-START transcription factor, conferred drought tolerance associated with improved root architecture (Yu et al. 2008). Among the nine drought-inducible *GmNAC* genes, with the exception of *GmNAC004* and *GmNAC012*, whose transcripts accumulated markedly in both root and shoot tissues, drought-inducible expression of *GmNAC003*, *015* and *020* genes significantly increased only in roots, whereas those of *GmNAC002*, *010*, *013* and *028* increased in shoots (Fig. 4). Collectively, these results indicate diverse functions of these stress-inducible GmNAC TFs, and also provide candidate *GmNAC* genes for generation of stress resistant transgenic plants possessing traits best suited for survival and/or decreased yield loss under stressed conditions.

It has been shown that the NAC TFs can regulate the drought stress response through both ABA-dependent and ABA-independent pathways (Fujita et al. 2004; Tran et al. 2004). Since expression of *GmNAC003* and *GmNAC004* is induced by both dehydration and ABA, these two proteins may be associated with drought response in an ABA-dependent manner (Fig. 5). Expression patterns of the remaining seven dehydration-inducible *GmNAC* genes in ABA-treated plants are similar to those of water control, suggesting that these GmNAC TFs may regulate expression of downstream genes independently of ABA. Except for *GmNAC20*, which is induced only by dehydration (Figs. 3, 5), transcripts of the other dehydration-inducible *GmNAC* genes showed significant increase under high salinity, indicating there indeed exists a cross-talk between drought and salt stress pathways as previously reported (Rabbani et al. 2003; Zhou et al. 2008). During our investigation we observed that the expression patterns of these nine *GmNAC* genes were also increased, to some extent, under water

control treatment. Similar phenomenon has also been observed by Rabbani et al. (2003) when they studied the expression profiles of abiotic stress-inducible genes in rice. This might be due to the abrasion/wounding, which occurred when the soybean plants were uprooted. Alternatively, we cannot rule out that this marginal induction of these *GmNAC* genes might be due to possibility of sudden change in growing conditions, as plants were first transferred from soil to excess water in order to remove the remaining soil from the roots, then allowed to grow hydroponically until frozen in liquid nitrogen. This hypothesis is supported by the fact that expression of most the genes, with the exception of *GmNAC10* and *GmNAC20*, returned to normal levels after 5 h of water control treatment (Fig. 5). Additionally, *GmNAC002*, *003*, *004*, *013*, *015* and *028* were also induced by cold stress (Fig. 5). Together, these results suggest the involvement of NAC TFs in the regulatory pathways of multiple abiotic stresses, implying their suitability as candidates for genetic manipulation aimed at enhancing tolerance to a variety of stresses.

Several *ANAC* and *ONAC* genes have been shown to be stress-responsive, including *ANAC002* (Lu et al. 2007), *ANAC019*, *ANAC055*, *ANAC072*, *ANAC081* (Fujita et al. 2004; Tran et al. 2004), *ANAC029* (Guo and Gan 2006), *ANAC092* (He et al. 2005), *SNAC1/ONAC46* (Hu et al. 2006) and *SNAC2/OsNAC6/ONAC48* (Nakashima et al. 2007; Hu et al. 2008). Among these ANAC and ONAC TFs, with the exception of ANAC029 and ANAC081, in planta evidence indicates that ANAC002 functions as a negative regulator, whereas ANAC019, ANAC055, ANAC072, SNAC1 and SNAC2/OsNAC6 function as positive regulators in the drought stress response (Tran et al. 2004; Hu et al. 2006, 2008; Lu et al. 2007; Nakashima et al. 2007). GmNAC002, GmNAC012 and GmNAC013 all share a high similarity along the entire protein sequence to ANAC002, suggesting that the GmNAC002, 012 and 013 TFs might function as negative regulators in stress signaling (Fig. 6). Therefore, for these genes, RNAi or loss-of-function mutagenesis would be appropriate biotechnological approaches to engineer improved drought tolerance. Phylogenetic analysis of the NAC domains retrieved from stress-responsive ANAC, ONAC and the isolated GmNAC proteins indicate that most of the stress-responsive ANAC, ONAC and GmNAC TFs cluster into the stress-responsive subgroup, with a few members randomly distributed among other subgroups (Fig. 7). This result suggests a close relationship among the stress-responsive NAC TFs across the plant kingdom, even though ONAC proteins originate from the monocotic rice plant while ANACs and GmNACs are exclusive to the dicotic *Arabidopsis* and soybean, respectively.

Overall, this report has provided the first insight into the previously uncharacterized GmNAC family of soybean, particularly in relation to stress responsiveness. These studies also provide useful information of potential stress-inducible candidate *GmNAC* genes and their associated stress-inducible promoters which can be exploited to engineer soybean plants for enhanced stress resistant. We are currently overexpressing the stress-responsive *GmNAC* genes in both *Arabidopsis* and soybean plants to further ascertain the significance of their function in planta.

Acknowledgment Funding supports from Missouri Soybean Merchandising Council (MSMC) to the soybean genomics drought research program at the University of Missouri are gratefully appreciated.

References

Aharoni A, Dixit S, Jetter R, Thoenes E, van Arkel G, Pereira A (2004) The SHINE clade of AP2 domain transcription factors activates wax biosynthesis, alters cuticle properties, and confers drought tolerance when overexpressed in *Arabidopsis*. Plant Cell 16:2463–2480

Aida M, Ishida T, Fukaki H, Fujisawa H, Tasaka M (1997) Genes involved in organ separation in Arabidopsis: an analysis of the cup-shaped cotyledon mutant. Plant Cell 9:841–857

Bhatnagar-Mathur P, Devi MJ, Reddy DS, Lavanya M, Vadez V, Serraj R, Yamaguchi-Shinozaki K, Sharma KK (2007) Stress-inducible expression of At DREB1A in transgenic peanut (*Arachis hypogaea* L) increases transpiration efficiency under water-limiting conditions. Plant Cell Rep 26:2071–2082

Bray EA (2004) Genes commonly regulated by water-deficit stress in *Arabidopsis* thaliana. J Exp Bot 55:2331–2341

Collinge M, Boller T (2001) Differential induction of two potato genes, *Stprx2* and *StNAC,* in response to infection by *Phytophthora infestans* and to wounding. Plant Mol Biol 46:521–529

Duval M, Hsieh TF, Kim SY, Thomas TL (2002) Molecular characterization of AtNAM: a member of the Arabidopsis NAC domain superfamily. Plant Mol Biol 50:237–248

Ernst HA, Olsen AN, Larsen S, Lo Leggio L (2004) Structure of the conserved domain of ANAC, a member of the NAC family of transcription factors. EMBO Rep 5:297–303

Fang Y, You J, Xie K, Xie W, Xiong L (2008) Systematic sequence analysis and identification of tissue-specific or stress-responsive genes of NAC transcription factor family in rice. Mol Genet Genomics 280:535–546

Fujita M, Fujita Y, Maruyama K, Seki M, Hiratsu K, Ohme-Takagi M, Tran L-SP, Yamaguchi-Shinozaki K, Shinozaki K (2004) A dehydration-induced NAC protein, RD26, is involved in a novel ABA-dependent stress-signaling pathway. Plant J 39:863–876

Fujita Y, Fujita M, Satoh R, Maruyama K, Parvez MM, Seki M, Hiratsu K, Ohme-Takagi M, Shinozaki K, Yamaguchi-Shinozaki K (2005) AREB1 is a transcription activator of novel ABRE-dependent ABA signaling that enhances drought stress tolerance in Arabidopsis. Plant Cell 17:3470–3488

Guo Y, Gan S (2006) AtNAP, a NAC family transcription factor, has an important role in leaf senescence. Plant J 46:601–612

Hazen SP, Wu Y, Kreps JA (2003) Gene expression profiling of plant responses to abiotic stress. Funct Integr Genomics 3:105–111

He XJ, Mu RL, Cao WH, Zhang ZG, Zhang JS, Chen SY (2005) AtNAC2, a transcription factor downstream of ethylene and auxin signaling pathways, is involved in salt stress response and lateral root development. Plant J 44:903–916

Hegedus D, Yu M, Baldwin D, Gruber M, Sharpe A, Parkin I, Whitwill S, Lydiate D (2003) Molecular characterization of *Brassica napus* NAC domain transcriptional activators induced in response to biotic and abiotic stress. Plant Mol Biol 53:383–397

Hu H, Dai M, Yao J, Xiao B, Li X, Zhang Q, Xiong L (2006) Overexpressing a NAM, ATAF, and CUC (NAC) transcription factor enhances drought resistance and salt tolerance in rice. Proc Natl Acad Sci USA 103:12987–12992

Hu H, You J, Fang Y, Zhu X, Qi Z, Xiong L (2008) Characterization of transcription factor gene SNAC2 conferring cold and salt tolerance in rice. Plant Mol Biol 67:169–181

James VA, Neibaur I, Altpeter F (2008) Stress inducible expression of the DREB1A transcription factor from xeric, *Hordeum spontaneum* L in turf and forage grass (Paspalum notatum Flugge) enhances abiotic stress tolerance. Transgenic Res 17:93–104

Jensen MK, Hagedorn PH, de Torres-Zabala M, Grant MR, Rung JH, Collinge DB, Lyngkjaer MF (2008) Transcriptional regulation by an NAC (NAM–ATAF1, 2-CUC2) transcription factor attenuates ABA signalling for efficient basal defence towards Blumeria graminis f sp hordei in Arabidopsis. Plant J 56:867–880

Jian B, Liu B, Bi Y, Hou W, Wu C, Han T (2008) Validation of internal control for gene expression study in soybean by quantitative real-time PCR. BMC Mol Biol 23:9–59

John I, Hackett R, Cooper W, Drake R, Farrell A, Grierson D (1997) Cloning and characterization of tomato leaf senescence-related cDNAs. Plant Mol Biol 33:641–651

Kacperska A (2004) Sensor types in signal transduction pathways in plant cells responding to abiotic stressors: do they depend on stress intensity? Physiol Plant 122:159–168

Kasuga M, Liu Q, Miura S, Yamaguchi-Shinozaki K, Shinozaki K (1999) Improving plant drought, salt, and freezing tolerance by gene transfer of a single stress-inducible transcription factor. Nat Biotechnol 17:287–291

Kim HS, Park OB, Yoo HJ, Jung MS, Lee MS, Han JH, Kim EK, Kim HS, Lim OC, Yun D-J, Chung SW (2007) Identification of a calmodulin-binding NAC protein (CBNAC) as a transcriptional repressor in *Arabidopsis*. J Biol Chem 282:36292–36302

Lu PL, Chen NZ, An R, Su Z, Qi BS, Ren F, Chen J, Wang XC (2007) A novel drought-inducible gene, ATAF1, encodes a NAC family protein that negatively regulates the expression of stress-responsive genes in *Arabidopsis*. Plant Mol Biol 63:289–305

Meng Q, Zhang C, Gai J, Yu D (2006) Molecular cloning, sequence characterization and tissue-specific expression of six NAC-like genes in soybean (*Glycine max* (L) Merr). J Plant Physiol 164:1002–1012

Mitsuda N, Seki M, Shinozaki K, Ohme-Takagi M (2005) The NAC transcription factors NST1 and NST2 of *Arabidopsis* regulate secondary wall thickenings and are required for anther dehiscence. Plant Cell 17:2993–3006

Mitsuda N, Iwase A, Yamamoto H, Yoshida M, Seki M, Shinozaki K, Ohme-Takagi M (2007) NAC transcription factors, NST1 and NST3 are key regulators of the formation of secondary walls in woody tissues of Arabidopsis. Plant Cell 19:270–280

Nakashima K, Tran L-SP, Nguyen VD, Fujita M, Maruyama K, Todaka D, Ito Y, Hayashi N, Shinozaki K, Yamaguchi-Shinozaki K (2007) Functional analysis of a NAC-type transcription factor OsNAC6 involved in abiotic and biotic stress-responsive gene expression in rice. Plant J 51:617–630

Oh SK, Lee S, Yu SH, Choi D (2005) Expression of a novel NAC domain–containing transcription factor (CaNAC1) is preferentially associated with incompatible interactions between chili pepper and pathogens. Planta 222:876–887

Olsen AN, Ernst HA, Leggio LL, Skriver K (2005) NAC transcription factors: structurally distinct, functionally diverse. Trends Plant Sci 10:79–87

Ooka H, Satoh K, Doi K, Nagata T, Otomo Y, Murakami K, Matsubara K, Osato N, Kawai J, Carninci P, Hayashizaki Y, Suzuki K, Kojima K, Takahara Y, Yamamoto K, Kikuchi S (2003) Comprehensive analysis of NAC family genes in *Oryza sativa* and *Arabidopsis thaliana*. DNA Res 20:239–247

Rabbani MA, Maruyama K, Abe H, Khan MA, Katsura K, Ito Y, Yoshiwara K, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2003) Monitoring expression profiles of rice genes under cold, drought, and high-salinity stresses and abscisic acid application using cDNA microarray and RNA gel-blot analyses. Plant Physiol 133:1755–1767

Riaño-Pachón DM, Ruzicic S, Dreyer I, Mueller-Roeber B (2007) PlnTFDB: an integrative plant transcription factor database. BMC Bioinformatics 8:42

Riechmann JL, Heard J, Martin G, Reuber L, Jiang C, Keddie J, Adam L, Pineda O, Ratcliffe OJ, Samaha RR, Creelman R, Pilgrim M, Broun P, Zhang JZ, Ghandehari D, Sherman BK, Yu G (2000) *Arabidopsis* transcription factors: genome-wide comparative analysis among eukaryotes. Science 290:2105–2110

Sablowski RWM, Meyerowitz EM (1998) A homolog of NO APICAL MERISTEM is an immediate target of the floral homeotic genes APETALA3/PISTILLATA. Cell 92:93–103

Sakuma Y, Maruyama K, Osakabe K, Qin F, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2006) Functional analysis of an Arabidopsis transcription factor, DREB2A, involved in drought-responsive gene expression. Plant Cell 18:1292–1309

Seki M, Kamei A, Yamaguchi-Shinozaki K, Shinozaki K (2003) Molecular responses to drought, salinity and frost: common and different paths for plant protection. Curr Opin Biotechnol 14:194–199

Sharp RE, Poroyko V, Hejlek LG, Spollen WG, Springer GK, Bohnert HJ, Nguyen HT (2004) Root growth maintenance during water deficits: physiology to functional genomics. J Exp Bot 55:2343–2351

Shoemaker RC, Schlueter J, Doyle JJ (2006) Paleopolyploidy and gene duplication in soybean and other legumes. Curr Opin Plant Biol 9:104–109

Simpson SD, Nakashima K, Narusaka Y, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2003) Two different novel *cis*-acting elements of *erd1*, a *clpA* homologous *Arabidopsis* gene, function in induction by dehydration stress and dark-induced senescence. Plant J 33:259–270

Souer E, van Houwelingen A, Kloos D, Mol J, Koes R (1996) The no apical meristem gene of petunia is required for pattern formation in embryos and flowers and is expressed at meristem and primordia boundaries. Cell 85:159–170

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 40. Mol Biol Evol 24:1596–1599

Thomashow MF (1999) Plant cold acclimation: freezing tolerance genes and regulatory mechanisms. Annu Rev Plant Physiol Plant Mol Biol 50:571–599

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876–4882

Tran L-SP, Nakashima K, Sakuma Y, Simpson SD, Fujita Y, Maruyama K, Fujita M, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2004) Isolation and functional analysis of *Arabidopsis* stress-inducible NAC transcription factors that bind to a drought-responsive *cis*-element in the *early responsive to dehydration stress 1* promoter. Plant Cell 16:2481–2498

Tran L-SP, Nakashima K, Sakuma Y, Osakabe Y, Qin F, Simpson SD, Maruyama K, Fujita Y, Shinozaki K, Yamaguchi-Shinozaki K (2007a) Co-expression of the stress-inducible zinc finger

homeodomain ZFHD1 and NAC transcription factors enhances expression of the *ERD1* gene in *Arabidopsis*. Plant J 49:46–63

Tran L-SP, Nakashima K, Shinozaki K, Yamaguchi-Shinozaki K (2007b) Plant gene networks in osmotic stress response: from genes to regulatory networks. Methods Enzymol 428:109–128

Tran L-SP, Urao T, Qin F, Maruyama K, Kakimoto T, Shinozaki K, Yamaguchi-Shinozaki K (2007c) Functional analysis of AHK1/ATHK1 and cytokinin receptor histidine kinases in response to abscisic acid, drought, and salt stress in *Arabidopsis*. Proc Natl Acad Sci USA 104:20623–20628

Uauy C, Distelfeld A, Fahima T, Blechl A, Dubcovsky J (2006) A NAC gene regulating senescence improves grain protein, zinc, and iron content in wheat. Science 314:1298–1301

Valliyodan B, Nguyen HT (2006) Understanding regulatory networks and engineering for enhanced drought tolerance in plants. Curr Opin Plant Biol 9:189–195

Vroemen CW, Mordhorst AP, Albrecht C, Kwaaitaal M, de Vries SC (2003) The CUP-SHAPED COTYLEDON3 gene is required for boundary and shoot meristem formation in *Arabidopsis*. Plant Cell 15:1563–1577

Wang W, Vinocur B, Altman A (2003) Plant responses to drought, salinity and extreme temperatures: towards genetic engineering for stress tolerance. Planta 218:1–14

Xie Q, Frugis G, Colgan D, Chua NH (2000) *Arabidopsis* NAC1 transduces auxin signal downstream of TIR1 to promote lateral root development. Genes Dev 14:3024–3036

Xie Q, Guo HS, Dallman G, Fang S, Weissman AM, Chua NH (2002) SINAT5 promotes ubiquitin-related degradation of NAC1 to attenuate auxin signals. Nature 419:167–170

Xiong Y, Liu T, Tian C, Sun S, Li J, Chen M (2005) Transcription factors in rice: a genome-wide comparative analysis between monocots and eudicots. Plant Mol Biol 59:191–203

Yamaguchi-Shinozaki K, Shinozaki K (2005) Organization of *cis*-acting regulatory elements in osmotic-and cold-stress-responsive promoters. Trends Plant Sci 10:88–94

Yamaguchi-Shinozaki K, Shinozaki K (2006) Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses. Annu Rev Plant Biol 57:781–803

Yoo YS, Kim Y, Kim YS, Lee JS, Ahn HJ (2007) Control of flowering time and cold response by a NAC-domain protein in *Arabidopsis*. PLoS ONE 2:e642

Yu H, Chen X, Hong YY, Wang Y, Xu P, Ke SD, Liu HY, Zhu JK, Oliver DJ, Xiang CB (2008) Activated expression of an Arabidopsis HD-START protein confers drought tolerance with improved root system and reduced stomatal density. Plant Cell 20:1134–1151

Zhang JZ, Creelman RA, Zhu JK (2004) From laboratory to field using information from Arabidopsis to engineer salt, cold, and drought tolerance in crops. Plant Physiol 135:615–621

Zhou QY, Tian AG, Zou HF, Xie ZM, Lei G, Huang J, Wang CM, Wang HW, Zhang JS, Chen SY (2008) Soybean WRKY-type transcription factor genes, GmWRKY13, GmWRKY21, and GmWRKY54, confer differential tolerance to abiotic stresses in transgenic Arabidopsis plants. Plant Biotechnol J 6:486–503

Zhu JK (2002) Salt and drought stress signal transduction in plants. Annu Rev Plant Physiol Plant Mol Biol 53:247–273