Isolation and characterization of GoRAV, a novel gene encoding a RAV-type protein in Galegae orientalis

Xiaofang Chen¹,², Zan Wang², Xuemin Wang², Jie Dong³
Jizhou Ren¹ and Hongwen Gao²*¹

¹Pratacultural college, GANSU Agricultural University, Lanzhou, Gansu 730070,
People’s Republic of China
²Institute of Animal Science, Chinese Academy of Agricultural Sciences,
Beijing, 100193, People’s Republic of China
³College of Animal Science and Technology, SHANXI Agricultural University,
Taigu, Shanxi 030801, People’s Republic of China

(Received 8 February 2009, accepted 25 March 2009)

RAV is a unique transcription factor in higher plants with AP2 and B3-like two distinct DNA-binding domains, but its roles in plant growth and development remains unknown. We have isolated a novel RAV family gene from Galegae orientalis, called GoRAV, which responds to cold induction. Sequence alignment showed that it shares high identity with other RAV family members in AP2 and B3 domain. Transient expression analysis using onion epidermal cells indicated that GoRAV protein is localized in the nucleus. Semi-quantitative RT-PCR (S-Q RT-PCR) analysis indicated that GoRAV is induced by cold, dehydration, high-salinity and abscisic acid, with the strongest induction in G. orientalis leaves during the early response to abiotic elicitors. GoRAV is more abundant in leaf than in stem, but is not expressed in root. This work adds a new member to the RAV family.

Key words: abiotic stress, Galegae orientalis, semi-quantitative RT-PCR, subcellular localization, transcription factor RAV

INTRODUCTION

Most cellular and developmental processes are controlled by gene expression that results from the mobilization of multiple different sets of transcription factors (Grandori et al., 2000; Stracke et al., 2001; Dimova et al., 2003; Köhler et al., 2003). The increasing numbers of genes that encode sequence-specific DNA-binding proteins or transcription factors are being cloned from higher plants (Meshi and Iwabuchi, 1995). Since physiological and biochemical changes during cold acclimation are revealed in the modulation of gene expression profiles, elucidating the transcriptional control mechanisms for stress response genes has been a productive area in plant stress-tolerance research. Expression of several transcription factors or DNA-binding proteins increases during cold treatment of Arabidopsis and other plants (Seki et al., 2001, 2002; Chen et al., 2002; Fowler and Thomashow, 2002; Sakuma et al., 2002). The AP2 domain was first identified as a DNA-binding domain conserved in a family of tobacco ethylene response element binding proteins (EREBPs) (Ohme-Takagi and Shinshi, 1995), and was later found to be conserved in Arabidopsis APETALA2 (AP2) (Jofuku et al., 1994), and other proteins (Elliott et al., 1996; Klucher et al., 1996; Wilson et al., 1996; Stockinger et al., 1997; Finkelstein et al., 1998). The B3 domains were identified as a highly conserved domain in VPI/ABI3 proteins, which are involved in abscisic acid (ABA) - regulated gene expression during seed maturation (McCarty, 1995). In addition to VPI/ABI3 proteins, B3-like domains are found in Arabidopsis ARFI (Ulmasov et al., 1997) and FUSS (Luerssen et al., 1998).

Plant transcription factors have been divided into two subfamilies based on the number and sequence of AP2/ERF domains present in the protein (Riechmann and Meyerowitz, 1998; Riechmann et al., 2000; Sakuma et al., 2002). The smaller subfamily contains transcription factors with two AP2/ERF domains, including APETALA2 (AP2) (Jofuku et al., 1994) and AINTEG-UMENTA (ANT) (Elliott et al., 1996) from Arabidopsis, and
Glossy15 (Moose and Sisco, 1994) from maize, all of which are involved in flower and seed development. The larger subfamily of transcription regulators contains only one AP2 domain, and these have been divided into several subgroups based on other sequence features (Sakuma et al., 2002).

A much smaller subgroup of AP2/ERF transcription regulators whose members all contain a single AP2/ERF followed by a B3 DNA-binding domain has also been identified and designated the RAV subfamily. Two of these proteins, RAV1 and RAV2, were isolated from *Arabidopsis*; the B3 domain in these proteins shows homology to the DNA-binding domain in VIVIPAROUS1 (VP1) from maize (McCarty et al., 1991) and its orthologue ABI3 in *Arabidopsis* (Giraudat et al., 1992). Both of these are involved in the regulation of gene expression in response to ABA. The AP2 and B3 domains of RAV1 bind independently to CAACA and CACCTG cis-elements respectively, and the presence of both domains increases the binding activity and specificity of the protein (Kagaya et al., 1999). It was revealed recently that RAV1 is a cold-responsive transcription factor that is likely to trigger another more effectual regulatory pathway than the well investigated DREB/CFB transcription factors do (Fowler and Thomashow, 2002).

The *Arabidopsis* genome contains several large transcription factor families with more than 100 members each, such as MADS, bHLH, MYB and AP2/EREBP (Riechmann and Ratcliffe, 2000; Bailey et al., 2003; Heim et al., 2003; Palenikova et al., 2003; Toledo-Ortiz et al., 2003; Jiang et al., 2004). AP2/EREBP genes are known to be involved in plant hormone signal transduction as well as response to biotic, pathogenic, and environmental stresses (Stockinger et al., 1997; Knight et al., 1999; Brown et al., 2003; Chakravarthy et al., 2003; Guterson and Reuber, 2004; Magome et al., 2004; Yi et al., 2004). A total of 144 AP2/EREBP members was found from the *Arabidopsis* genome and was divided into five groups: AP2, RAV, DREB, ERF and others (Sakuma et al., 2002). Recent publication indicated that the family has 145 members and it is no longer plant-specific since homologs have been revealed from the cyanobacterium *Trichodesmium erythraeum*, the ciliate *Tetrahymena thermophila*, and the viruses *Enterobacteria phage Rb49* and *Bacteriophage Felix 01* (Magnani et al., 2004).

Leguminous plants are of great importance worldwide, especially in countries with Mediterranean climates like Tunisia. They provide human food, forage, wood, industrial and medicinal products. They contribute to soil fertility and prevent soil erosion. They are good colonizers of poor soils under extreme climatic conditions, due in large part to nitrogen-fixing symbiosis with Legume Nodulating Bacteria (LNB) belonging to the α and β subclasses of Proteobacteria (Chen et al., 2001; Moulin et al., 2001; Zakhia and de Lajudie, 2001). *G. orientalis* is a promising perennial nitrogen-fixing forage legume used in Finnish agriculture for forage production and soil improvement (Varis, 1986). *G. orientalis* has a strong root system and spreads effectively by underground stolons (Lindström, 1989).

In this study, we isolated and characterized the *GomeraV* in response to abiotic stresses. As far as we are aware, this work represents the first published RAVs study of *G. orientalis*.

**MATERIALS AND METHODS**

**Plant materials and stress treatments** Mature seeds of *G. orientalis* were surface-sterilized and germinated in Petri dishes at 24°C in 12 hr light/12 hr dark. After 2–3 days, they were transplanted into pots with 3:1 (v/v) vermiculite with perlite and grown in a growth chamber for 5 weeks. The stress treatments were performed as follows: for salt, dehydration, and ABA treatments, the plants were gently pulled out of the pots and cleaned with tap water. We placed the roots of seedlings into a beaker containing solutions of 800 ml 250 mM NaCl, 20% PEG, 100 μM ABA for 2 hr, 4 hr, 8 hr, 12 hr or 24 hr. For cold treatment, the plants were placed in a 4°C refrigerator for 2 hr, 4 hr, 8 hr, 12 hr or 24 hr. The stress-treated and control seedlings were harvested at various times, then quickly frozen in liquid nitrogen and stored at −80°C for use.

**Isolation of *G. orientalis* total RNA** Total RNA was isolated from leaf samples with Trizol® Reagents (Invitrogen, USA) according to the manufacturer’s instructions. Total RNA concentration was determined spectrophotometrically and samples stored at −80°C.

**5’- and 3’-Rapid amplification of complementary DNA ends of *G. orientalis* RAV** Total RNA was extracted from the seedlings treated by cold for 8 hr with TRIZOL® Reagent (Invitrogen, USA) according to the manufacturer’s protocol. cDNA was synthesized from the total RNA by the RevertAid™ First Strand cDNA Synthesis Kit (MBI) following the manufacturer’s instructions. Two degenerate primers (AP2-1: 5’- AGGCCTTGCGCTNGNACNTTY -3’; AP2-2: 5’- TCCAGGAAGTTGAGACG -3’) were designed using the AP2 DNA-binding domain of previously cloned AP2/EREBP transcription factors. The AP2 DNA-binding domain was amplified by polymerase chain reaction (PCR) under the following conditions: 94°C 30 sec, 72°C 2 min for 5 cycles; 94°C 30 sec, 70°C 30 sec, 72°C 2 min for 5 cycles; 94°C 30 sec, 68°C 30 sec, 72°C 2 min for 36 cycles; The PCR product was purified and cloned into pMD18-T vector (Takara) and sequenced. Primer GSP1 (5’- CTTTGCTTGGCGAGTCAACTC -3’) was designed using the AP2 region cDNA sequence. Generation of the first strand cDNA and the 3’-RACE
PCR were performed following the 3'-Full RACE Core Set Ver. 2.0 (Takara) manufacturer’s instructions. The PCR program was as follows: 94°C for 3 min, followed by 25 cycles of 94°C for 30 sec, 58°C for 30 sec, 72°C for 2 min, and extension at 72°C for 10 min. The PCR product was purified and cloned into pMD18-T vector and sequenced.

Amplification of the 5'-cDNA end sequence was performed by Smart™ RACE cDNA amplification kit (Clontech, USA). GSP2 (5'- GCCAGCGACAGCACCACGGA AATAATTTTGCA -3') was designed using the 3' end of the cDNA sequence. The PCR program was as follows: 94°C 30 sec, 72°C 2 min for 5 cycles; 94°C 30 sec, 70°C 30 sec, 72°C 2 min for 5 cycles; 94°C 30 sec, 68°C 30 sec, 72°C 2 min for 25 cycles; and extension at 72°C for 10 min. The PCR product was purified and cloned into pMD18-T vector and sequenced.

Isolation of the full-length coding sequence encoding GoRAV from G. orientalis The 3' and 5' sequences obtained from RACE were assembled and aligned to determine full-length sequences. Primers GSP3 (5'-GAAAAATGGGAAGGATGGTGTC-3') and GSP4 (5'-TTATAATGTGCTCAATTTGACC-3') were designed to amplify the full-length GoRAV cDNA and GoRAV DNA, spanning the ORF. For isolation of the GoRAV genomic clone, total genomic DNA was extracted from G. orientalis leaves using CTAB protocol (Dellaporta et al., 1983). Genomic DNA was used as a template and the annealing temperature for PCR was reduced to 58°C.

Bioinformatics analysis The full-length cDNA sequences and the deduced amino acid sequences were compared to previously characterized DNA and protein sequences in GenBank by the software DNAMAN. The nuclear localization signal (NLS) of GoRAV was analyzed by Predict Protein (http://www.predictprotein.org/).

Semi-quantitative RT-PCR analysis First-strand cDNA was prepared from 5-week-old G. orientalis seedlings subjected to one of the following conditions: cold treatment at 4°C, high-salt treatment with 250 mM NaCl, ABA treatment with 100 μM ABA or dehydration with 20% PEG. Total RNA concentrations were estimated at 260 nm with a spectrophotometer, and 5 μg of the RNA was reverse-transcribed in a total volume of 20 μl using a RevertAid™ First Strand cDNA Synthesis Kit (MBI). Primers GSP3 and GSP4 were used in PCR amplification. Different numbers of cycles were analyzed to establish a point at which no band saturation occurred. S-Q RT-PCR was operated under the following conditions: 3 min at 94°C, followed by 25 cycles of 30 sec at 94°C, 30 sec at 56°C, 1 min at 72°C, and a final extension at 72°C for 10 min. PCR reaction of the actin gene was performed as an internal control, verifying that a similar quantity of first strand cDNA was used for each RT-PCR reaction. Each RT-PCR reaction was performed in triplicate.

Subcellular localization of GoRAV protein Since the RAV domain contains two DNA-binding domains and the deduced RAV protein contains two stretches of basic residues that could act as NLS, we studied the subcellular localization of the RAV protein. A GoRAV-green fluorescent protein (GFP) gene fusion was constructed and transformed into onion epidermal cells to express the fusion protein. To amplify the coding sequence of GoRAV, two primers were designed, all with Spe I restriction sites (RAVGFP1: 5'- CGGACTAGTAGGAA GGAGGATGTTGACAT-3', RAVGFP2: 5'-CGGACTAGTTAATGCTC AAATATCATTGCTT-3'). The PCR product was digested with Spe I, and ligated into the pCAMBIA1302 vector digested with the same restriction enzyme. The fusion was then transformed into onion epidermal cells using a gene gun. Subcellular localization of transiently expressed pCAMBIA1302-GoRAV fusion was detected by confocal microscopy.

RESULTS

Isolation of the Full-Length cDNA of G. orientalis RAV Based on the alignment of consensus sequences, primers AP2-1 and AP2-2 were designed to cover a region that showed a high degree of similarity among the different sequences. Using these primers, a 456 bp cDNA fragment was isolated from G. orientalis leaves after treatment at 4°C for 8 hr. To obtain a full length cDNA, 5' and 3' RACE were employed to extend both ends of the putative RAV homolog. A 1040 bp fragment was generated with 3'RACE, and a 1012 bp fragment containing a 475 bp overlapping sequence with the 3'-cDNA end sequence was obtained by 5'RACE. From these, a 1626 bp full-length cDNA sequence was obtained. Genomic PCR using primers GSP3 and GSP4, designed from the 5' and 3' UTRs, and subsequent sequencing of the PCR product revealed that the gene has no introns. Sequence alignment showed that the isolated gene is similar to the RAV genes reported in other species, and phylogenetic analysis revealed that the gene is a RAV-like gene. We designated the gene GoRAV, and deposited the complete sequence in GenBank (Accession: FJ223569).

Sequence analysis of G. orientalis RAV Nucleotide sequence analysis revealed that the cold-inducible gene contains an open reading frame of 1164 bp and encoded a novel protein of 387 amino acids with a calculated molecular mass of 41.954 kDa and a pI of 8.538. The unique features of GoRAV led us to further characterize the molecular mechanisms involved in plant cold-resistance. Sequence analysis revealed that GoRAV contains two distinct DNA-binding domains uniquely found
Fig. 1. Nucleotide and deduced amino acid sequences of *G. orientalis* RAV gene. The amino acid sequences in boxes represent the conserved AP2 and B3 DNA-binding domains that are characteristic of RAV transcription factors. The locations of the conserved YRG and RAYD elements are indicated in brackets. The underlined amino acid sequence is a WAAEIRD box that is conserved in proteins with one AP2 domain. The amino acids that potentially act as a nuclear localization signal are shown in bold and italics, and an acidic N-terminal region that may act as a transcriptional activation domain is in bold.
Isolation and characterization of \textit{GoRAV} in higher plants, AP2 and B3, which are characteristic of other RAV DNA-binding proteins (Kagaya et al., 1999). The AP2 domain, composed of 56 amino acids, and present in the N-terminal region of \textit{GoRAV}, has been extensively characterized in various plant proteins, including APETALA2 and ERFs. The AP2 domain of \textit{GoRAV} contains conserved YRG and RAYD elements, which are predicted to play an important functional role. \textit{GoRAV} also possesses a conserved 7-amino acid sequence motif, referred to as the WAAEIRD box, within the AP2 domain (Okamuro et al., 1997). In addition to the AP2 domain in the predicted \textit{GoRAV} gene product, 119 amino acids show significant sequence similarity to the B3 domains of \textit{Arabidopsis} (\textit{Arabidopsis thaliana}) and \textit{VP1} (maize). \textit{GoRAV} contains two nuclear localization sequences in the C-terminal region and the amino acids

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig2}
\caption{Comparison of amino acid sequences of the \textit{G. orientalis} RAV with other reported RAV proteins. A: Multiple sequence alignments of the predicted amino acid sequence of \textit{GoRAV} with proteins in GenBank with the highest homology. The AP2 domain is underlined, and the B3 domain is indicated with an arrowhead. The amino acid sequences of \textit{Arabidopsis thaliana} RAV1 (accession no. NM_101197) and RAV2 (accession no. AY091069), \textit{Oryza sativa} RAV2 (accession no. NM_192160), \textit{G. max} RAV (accession no. DQ147914), \textit{Capsicum annuum} RAV (accession no. AT27830) and \textit{Nicotiana tabacum} RAV (accession no. EU870518) were aligned. B: Phylogenetic analysis of the deduced amino acid sequence of the \textit{GoRAV} gene. The phylogenetic tree shows the relationships between \textit{GoRAV} and other RAV proteins.}
\end{figure}
required for transcriptional activation in the N-terminal region (Fig. 1).

Comparisons of the amino acid sequences between the GoRAV and other RAV genes are presented in Fig. 2A. The predicted GoRAV gene product shares 55.13% identity with the Arabidopsis RAV1 protein (accession no. NM_101197), 57.61% identity with the Arabidopsis RAV2 protein (accession no. AY091069) and 72.24% identity with the Glycine max RAV-like DNA-binding protein (accession no. DQ147914). These proteins share very high homology in the DNA-binding domains, and RAV transcription factors from G. max show high-level homology not only in the AP2 and B3 domains but also outside of the conserved domains.

A phylogenetic tree was constructed to analyze the relationship between the protein encoded by GoRAV and other RAV proteins obtained from GenBank (Fig. 2B). The phylogenetic tree revealed a clear separation between RAV proteins from monocotyledonous and dicotyledonous species, and branch distribution showed that GoRAV had higher homology to RAV proteins from dicots, and lower homology to the monocot. Furthermore, it had the highest homology with the RAV protein from G. max because they both belong to the Leguminosae.

Expression profiles of *G. orientalis* RAV in response to abiotic stresses and in different tissues To examine the spatial expression pattern of the GoRAV gene, we monitored the level of expression in different *G. orientalis* vegetative tissues, including roots, stems and leaves, by S-Q RT-PCR analysis (Fig. 3). The accumulation of GoRAV RNA was more abundant in stems than in leaves, and GoRAV was not expressed in roots. The results showed strict control of the GoRAV gene in different organs.

GoRAV expression was up-regulated within 4 hr after exposure to PEG treatment, and peaked at 12 hr. At 24 hr, transcripts declined and were almost at the level of 8 hr transcripts. When the plants were treated with NaCl, there was slight difference in transcript levels throughout the time course. Under ABA stress, the amount of transcript increased slowly, reaching a peak at 24 hr. The amount of GoRAV transcript increases steadily after cold treatment, reaching a peak at 24 hr. This result implied that the GoRAV might be involved in the ABA-dependent signalling pathway (Fig. 4). Taken together, the results suggested that GoRAV is not only involved in drought and low temperature stress signal pathways, but also influences ABA signal transduction at the transcriptional level.

Subcellular localization of *G. orientalis* RAV To examine if *G. orientalis* RAV was localized in the nucleus, a pCAMBIA1302-GoRAV gene fusion was constructed. Both the recombinant plasmid and a control pCAMBIA1302 vector were transformed into onion epidermal cells using a gene gun. When the cultured onion epidermal cells producing pCAMBIA1302-GoRAV fusion
proteins were examined by confocal microscopy, a strong fluorescence signal was observed only in the nucleolus (Fig. 5). In contrast, the GFP signal was distributed throughout the onion epidermal cells with the control pCAMBIA1302 vector. These results indicated that *G. orientalis* RAV is a nuclear protein that possibly functions as a transcription factor.

**DISCUSSION**

A relatively small family of RAV proteins has been demonstrated to belong to the AP2/EREBP family (Riechmann et al., 2000). Using RACE, we isolated and characterized the gene for a novel RAV protein (GoRAV) from *G. orientalis* leaves treated with cold for 8 hr.

As shown in Fig. 2A, GoRAV contains highly conserved AP2 and B3 DNA-binding domains, which are characteristic of other transcription factors, including ERFs, DREBs, VP1 and ABI3 (Suzuki et al., 1997; Ezcurra et al., 2000; Ohta et al., 2000; Sakuma et al., 2002). Extensive studies have revealed that AP2 DNA-binding domain plays a critical role in the specific binding of AP2 domain proteins to their target sequence, known as GCC box, and to dehydration responsive elements (DRE) and C-repeats (CRT) (Ohme-Takagi and Shinshi, 1995; Hao et al., 1998; Sakuma et al., 2002).

The other DNA-binding domain present in GoRAV is the B3 domain, conserved in the ABI3/VP1 family of regulatory proteins (Suzuki et al., 1997; Luerssen et al., 1998). The B3 domain, consisting of approximately 120 amino acids, has been demonstrated in many proteins, such as ABI3/VP1, FUSCA3, ARF1 (AUXIN RESPONSE FACTOR1), LEC2 and RAV (Suzuki et al., 1997; Ulmasov et al., 1997; Luerssen et al., 1998; Kagaya et al., 1999; Stone et al., 2001). To date, the direct evidence has showed that B3 domain responsible for DNA-binding activity has not been found, with the exception of ABI3/VP1 and RAV. In one protein, RAV1, both the AP2 and B3 domains have been demonstrated to cooperatively bind to their specific target sequences (CAACA and CACCTG) (Kagaya et al., 1999). However, it is unclear if the CAACA and CACCTG motifs are the cis-regulatory elements existing in the promoter regions of target genes that are transcriptionally regulated by RAV1. The core sequences of the AP2 domain in GoRAV were 23%–50% identical to those of other AP2 family proteins (data not shown), whereas 74%–88% identity was shown between the GoRAV and the RAV proteins from other plants. This suggests significant differences between the AP2 family and the RAV proteins in the amino acid sequences of their AP2 domains. As in the AP2 domain, a relatively high level of identity (65%–82%) was shown between the B3 domains of the GoRAV and RAV proteins from other plants, whereas B3 domain-containing proteins, such as ARFs and VP1, were 23%–32% identical to GoRAV. The relatively low levels of conservation within the two DNA-binding domains of GoRAV may affect the structure of GoRAV, possibly resulting in binding to a specific target gene.

Plant transcription factors contain NLS that vary in sequence, organization and number (Liu et al., 1999). As expected for transcription factors, the pCAMBIA1302-GoRAV fusion proteins were localized in the nucleus of onion epidermal cells under normal conditions, and GoRAV contains two NLS in the C-terminal region.

In this study, GoRAV expression was induced by exogenous ABA, drought, low-temperature and high-salinity, which showed that it is involved in the ABA-dependent signal transduction pathway in *G. orientalis*. *GmRAV* mRNA abundance increased following leaf treatment with ABA, and decreased following BR treatment. *GmRAV* transcript abundance in soybean leaves was enhanced during dark- and ABA-induced senescence that caused a decrease in chlorophyll content (Zhao et al., 2008). The expression of the *CaRAV1* gene was strongly induced in pepper leaves during early pathogen infection,
abiotic elicitors and environmental stress (Kee et al., 2006).

Since the expression of GoRAV is significantly higher in G. orientalis under cold and PEG treatments, GoRAV may contribute to cold hardiness. Since GoRAV encodes a protein with highly conserved DNA-binding domains, it is likely to be involved in transcriptional regulation of other cold-induced genes in G. orientalis. Therefore, not only GoRAV, but also downstream genes potentially activated by these proteins could contribute to cold hardiness. Although this study establishes the involvement of GoRAV in environmental-stress, further studies are needed to determine the actual functions and involvement of these genes in environmental-stress hardiness in G. orientalis.

This research was financially supported by the National Basic Research Programme of China (973 Programme) (2007CB108902) and the National Key Technology R&D Program (2008BAD3B01).

REFERENCES


X. CHEN et al.


