Mol. Cells *31*, 55-64, January 31, 2011
DOI/10.1007/s10059-011-0001-2

Molecules and Cells

Phylogenetic and Expression Analysis of RNA-binding Proteins with Triple RNA Recognition Motifs in Plants

Lila Peal[1], Niranjani Jambunathan[1,2], and Ramamurthy Mahalingam[1,*]

The superfamily of RNA binding proteins (RBPs) is vastly expanded in plants compared to other eukaryotes. A subfamily of RBPs that contain three RNA recognition motifs (RRMs) from the Arabidopsis (24), rice (19) and poplar (37) genomes was analyzed in this study. Phylogenetic analysis with full-length protein sequences of 80 RBPs identified nine clades. The largest clade, comprising 23 members, showed high homology to human RBPs involved in oxidative signaling. Digital northern analysis revealed that Arabidopsis RBPs are transcriptionally responsive to biotic, abiotic and hormonal treatments. Northern blot analysis of eight Arabidopsis RBPs belonging to the tobacco RBP45/47 family showed that these genes respond to ozone stress. AtRBP45b, which shows closest homology to the yeast oxidative stress regulatory protein, CSX1, was expressed in multiple tissues. Two novel splice variant forms of AtRBP45b were identified by 3'RACE analysis. Based on RT-PCR, splice variant AtRBP45b-SV1 was observed only in response to mechanical wounding caused by pathogen or chemical infiltrations and was not detectable in response to salt or temperature stress. Electrophoretic mobility shift assay demonstrated that recombinant full-length and splice variant forms of AtRBP45b bound synthetic RNA. Identifying *in vivo* RNA targets of AtRBP45b will aid in determining the precise functional role of these proteins during oxidative signaling.

INTRODUCTION

Post-transcriptional control of gene expression plays a vital role in the normal growth and development of eukaryotes. This level of regulation includes mRNA stability, capping, splicing, alternative splicing, polyadenylation, transport and translation (Anantharaman et al., 2002). This regulation is achieved either directly by RNA-binding proteins (RBPs) or indirectly via the interactions between RBPs and other regulatory factors (Fedoroff, 2002; Lorkovic and Barta, 2002).

In the genome of the model plant *Arabidopsis thaliana*, the superfamily of RBPs is largely expanded. There are 279 RBPs in Arabidopsis (Anantharaman et al., 2002), as compared to 100 in *Caenorhabditis elegans* (Consortium, 1998) and 117 in Drosophila (Lasko, 2000). However, the functions of only a few plant RBPs have been reported thus far (Lorkovic, 2009). Light-regulated translation of chloroplastic mRNAs is controlled post-transcriptionally utilizing RBPs encoded in the nuclear genome (Danon and Mayfield, 1991; Jensen et al., 1986; Mussgnug et al., 2005; Rochaix, 2001). Arabidopsis mutant studies have shown that a battery of different RBPs is involved in the post-transcriptional regulation of flowering (Lim et al., 2004; Macknight et al., 1997; Schomburg et al., 2001). The role of plant RBPs in hormone signaling has gained increasing attention. Several genes in the abscisic acid (ABA) signaling pathway have been identified as RBPs (Hugouvieux et al., 2001; Lu and Fedoroff, 2000). RBPs have been shown to interact with single-stranded telomeric DNA and to repress the telomerase activity (Kwon and Chung, 2004). Further, plant RBPs interact with DEAD box RNA helicases and fibrillarin to form a ribonucleoprotein complex involved in ribosomal RNA metabolism (Gendra et al., 2004). Recently, it was shown that the bacterial type III effector protein HopU1 mono-ADP ribosylates plant RBPs, which leads to the quelling of the plant innate immune system and the rendering of the plants susceptible to pathogens (Fu et al., 2007). Three RBPs belonging to the UBA2 group were shown to be involved in a novel wound signal transduction pathway that is independent of jasmonic acid (Bove et al., 2008). A zinc finger-containing, glycine-rich RNA binding protein, atRZ-1a, was shown to be important during seed germination and seedling development, especially under drought and salt stress conditions (Kim et al., 2007). The glycine-rich RNA binding proteins (GRPs) AtGRP7 and AtGRP8 were rapidly upregulated in response to peroxide-induced oxidative stress (Schmidt et al., 2009). AtGRP7 is abundantly expressed in guard cells and is important for regulating stomatal movement during dehydration and salt stress (Kim et al., 2008). GRPs have been shown to function as RNA chaperones during adaptation to cold stress in both monocots and dicots (Kim et al., 2009; 2010). These studies exemplify the myriad functional roles of plant RBPs during normal plant development and in response to environmental perturbations.

Classification of the RBPs has been based on the various conserved domains found in the protein sequences. These

[1]Department of Biochemistry and Molecular Biology, Oklahoma State University, Stillwater, USA, [2]Present Address: Monsanto, Saint Louis, USA
*Correspondence: ramamurthy.mahalingam@okstate.edu

Received August 11, 2010; revised September 17, 2010; accepted September 29, 2010; published online November 25, 2010

Keywords: AtRBP45, electrophoretic mobility shift assay, oxidative signaling, RNA binding proteins, splice variant, stress

domains include RNA recognition motifs (RRMs), serine-arginine-rich motifs (Manley and Tacke, 1996), RGG boxes (Kiledjian and Dreyfuss, 1992), K homology motifs (Burd and Dreyfuss, 1994), and chloroplast RNA splicing and ribosome maturation (CRM) domains (Barkan et al., 2007). RBPs containing RRMs are the most common among eukaryotes (Anantharaman et al., 2002; Burd and Dreyfuss, 1994). The RRM consists of an octomer sequence, RNP1, and a hexamer sequence, RNP2, both of which are usually contained within an approximately 80 amino acid region (Burd and Dreyfuss, 1994). The RRMs show conservation at the structural level rather than at the sequence level (Lorkovic and Barta, 2002).

In this study, phylogenetic analysis of Arabidopsis, rice and poplar RBP protein sequences that contain three RRMs was conducted. Various informatic tools were used to ratify the phylogenetic groupings and to predict possible functions of these proteins. In order to begin assigning functional roles to the Arabidopsis RBPs, steady state levels of a subset of eight Arabidopsis genes were analyzed during ozone-induced oxidative stress. One of the genes, *AtRBP45b*, has high homology to the yeast CSX1 protein, a global regulator of oxidative stress (Rodriguez-Gabriel et al., 2003). Expression patterns of *AtRBP45b* and its splice variants were analyzed in response to oxidants and abiotic and biotic stresses. Electrophoretic mobility shift assays demonstrated that recombinantly expressed full-length and splice variant forms of AtRBP45b protein bound U-rich RNA.

MATERIALS AND METHODS

Database analysis to identify triple RRM motif-containing proteins in plants

A search for three RRM-containing proteins in Arabidopsis and rice was conducted using Putative Orthologous Groups/Plant RBP database (http://plantrbp.uoregon.edu) (Walker et al., 2006) and the poplar genome database (http://genome.jgi-psf.org/Poptr1_1/Poptr1_1.home.html). The sequences were analyzed using the NCBI's Conserved Domain Search Database (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi) (Marchler-Bauer et al., 2005). The rice sequences were retrieved from the POG database and further confirmed through the Genome Cluster database (http://bioinfo.ucr.edu/projects/GCD).

Sequence alignment and phylogenetic tree construction

The complete predicted protein coding sequences of the RBPs containing three RRMs from Arabidopsis (24), Rice (19) and Poplar (37) were used for multiple sequence alignments. The evolutionary history was inferred using the Neighbor-Joining method (Saitou and Nei, 1987). The bootstrap consensus tree inferred from 5000 replicates was used to represent the evolutionary history of the taxa analyzed. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates were collapsed. The tree was drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method (Zuckerkandl and Pauling, 1965) and are in the units of *number of amino acid substitutions per site*. All positions containing gaps and missing data were eliminated from the dataset (Complete Deletion option). Phylogenetic analyses were conducted in MEGA4 (Tamura et al., 2007).

Motif identification

To further analyze these protein sequences, the Multiple Em for Motif Elicitation (MEME) (Bailey et al., 2006) was used. The parameters were set to identify a maximum of five motifs (6-80 amino acids) with any number of repetitions per sequence. The motifs were set to be present in at least 10 sites among the 80 RBPs. The results from MEME analysis were exported to Motif Alignment and Search Tool (MAST) (Bailey and Gribskov, 1998) to identify proteins in the PubMed database that have similar organization of these five motifs found in plant RBPs with triple RRMs.

Genevestigator analysis

The stress response profiles of the 24 Arabidopsis RBPs were retrieved using the Genevestigator program (Zimmermann et al., 2005). The locus identifiers of the 24 Arabidopsis genes were entered into this program. Three of the loci, At1g47490, At2g43410 and At1g60900, are not represented in the Affymetrix ATH1 gene chip. Along with the stress response profile, the developmental profile and the anatomy profile data for these genes were retrieved. All available ATH1 gene chip data (3110 experiments) were included in this analysis.

Plant materials

Seeds of *Arabidopsis thaliana* ecotype Wassilewskija (Ws-0) or Columbia (Col-0) were incubated at 4°C for three days before sowing in pots containing Metro-Mix 200 potting mix (Scotts-Sierra Horticultural Company, USA). Plants were grown in growth chambers with 10 h light (100 μmoles m^{-2} s^{-1}) and 14 h of darkness. The temperature in the growth chamber was 20 $\pm$ 2°C. Plants were irrigated with 15-5-5 Peter's Excel fertilizer (Scotts-Sierra Horticultural Company, USA; 1/8 tablespoon/gallon of water) every 3-4 days. Plants were grown for 4-5 weeks in these growth chambers prior to stress treatments.

Stress treatments

Ozone treatment: The ozone generation set-up has been described (Mahalingam et al., 2006). Ozone-sensitive Ws-0 ecotype plants were exposed to 290 $\pm$ 20 nl L^{-1} of ozone for six hours. Entire rosettes were collected 1 and 4 h after initiation of the ozone treatment and 2, 6, and 18 h after the end of treatment. Control plants maintained under ambient ozone were also collected for each time point.

Hydrogen peroxide, glucose/glucose oxidase and xanthine/xanthine oxidase treatment: Approximately 15 leaves of five-week-old, WT Col-0 plants were infiltrated with either 10 mM H_2O_2 or 2.5 mM glucose and 2.5 U/ml Glucose Oxidase or 0.5 mM Xanthine and 0.5 U/ml Xanthine Oxidase using a needleless syringe. Control leaves were infiltrated with H_2O or glucose or xanthine, respectively. Infiltrated leaves were collected 1, 8, and 24 h after treatment initiation.

Salt, heat and cold stress: Five-week-old, WT Col-0 plants were treated with 0.5 M NaCl for 24 h while control plants were treated with water. Plants were incubated at 40°C for three hours or at 4°C for 24 h for heat and cold stress treatments, respectively. Control plants were maintained at normal growth conditions. Entire rosettes were collected 1, 8, and 24 h after start of treatment.

Bacterial pathogen infection: approximately 15 leaves of five-week-old, WT Col-0 plants were infiltrated with virulent *Pseudomonas syringae* strain DC3000 or avirulent *P. syringae* DC3000 containing *AvrRpt2*. Approximately 1×10^5 colony forming units were resuspended in 10 mM $MgCl_2$ and used for infiltrations. Control leaves were infiltrated with 10 mM $MgCl_2$ buffer. Infiltrated leaves were collected 1, 2, 4, and 6 h after infiltration.

RNA isolations, cDNA synthesis and RT-PCR

Total RNA was isolated from control and stressed plant materi-

als using the RNeasy Plant Mini kit (Qiagen, USA) following the instructions in the manual. cDNA synthesis was carried out as described earlier (Mahalingam et al., 2005).

Primers were designed to amplify the complete, predicted open reading frame of each of the eight AtRBP genes belonging to the tobacco RBP45/47 family (Supplementary Table S1). PCR was conducted using cDNA as template on a PCR MasterCycler (Eppendorf, USA). Amplification pro-ducts were analyzed on a 1% Tris-Acetic acid-EDTA gel. Discrete amplicons of the expected sizes were ligated into pGEM-T Easy vector (Promega, USA) following the manufac-turer's protocol. Colony PCR products from positive amplifications of the expected size were used for generating probes for northern hybridizations.

In order to identify both the splice variant forms and the full-length AtRBP45b transcripts in a single reaction, 3 μl of cDNA was used as template. PCR reactions were conducted using Takara Taq Polymerase in the following reaction: 10 μl water, 2 μl 10× PCR Buffer, 1.2 μl 25 mM $MgCl_2$, 1 μl 5 μM forward primer designed from exon 4 (5′-gtggattagatgcatctgtcacggatg-3′), 1 μl 5 μM reverse primer designed from the common 3′ UTR region (5′-cctgtggtgtcagctggtaa-3′), 3 μl cDNA, and 0.8 μl Takara Taq. PCR conditions were as follows: initial denaturation at 94°C for 2 min, 30 cycles of 94°C for 30 s, 60°C for 30 s, and 72°C for 1 min, and a final extension at 72°C for 7 min. Products were run on a 1.5% agarose gel cast in TAE for 30 min at 150 V.

In order to identify the differential expression patterns of the full-length *AtRBP45b,* PCR was repeated with 1 μl of cDNA with the number of PCR cycles reduced to 20. The cDNA synthesis and RT-PCR experiments were repeated twice from two independent biological replications of the stress treatments.

Northern blots and hybridizations

Ten micrograms of RNA was used for making Northern blots. Details of northern blotting procedures and hybridizations were as described earlier (Mahalingam et al., 2005).

3′ RACE for identifying splice variants of *AtRBP45b*

3′ RACE was done as described in the FirstChoice RLM-RACE Kit (Ambion) with a few modifications. For cDNA synthesis, 1 μg total RNA, 4 μl dNTP mix, 2 μl 3′RACE Adapter, 4 μl 5× RT Buffer, 1 μl RNAse Inhibitor, 1 μl Superscript III Reverse Transcriptase (Invitrogen), 2 μl 0.1M DTT, and nuclease free water to a total volume of 20 μl were combined in a microfuge tube and incubated for 1 h at 42°C. Primary RACE PCR was set up with 1 μl of cDNA, 2.5 μl 10× PCR Buffer, 2 μl dNTP mix, 1.5 μl $MgCl_2$, 1 μl 3′RACE Outer Primer (10 μM 5′-gcgagcacagaattaa tacgact-3′), 1 μl 3′ RACE gene specific primer (10 μM 5′-atattacggaggctactctggtggag-3′ (FL) and 5′-gtggattagatgcatctgtc acggatg-3′, 1 μl Takara Taq, and nuclease free water to a total volume of 25 μl in a microfuge tube. PCR was performed using the following parameters: initial denaturation for 5 min at 94°C, followed by 30 cycles of amplification (94°C 30 s, 63°C 30 s, 72°C 30 s), and final extension at 72°C for 7 min. A secondary PCR for the splice variant was performed using 1 μl of primary PCR reaction. All conditions were the same except a 3′ RACE gene specific internal primer 5′-tcgttcagttttccgagaagtaa-3′ was used. The reaction products were then cloned into the Promega pGEM-T Easy Vector, colonies were screened by PCR, and plasmids bearing the right-sized inserts were sequenced.

Isolating recombinant AtRBP45b proteins

The full-length *AtRBP45b* coding sequence was amplified using gene specific primers with *Kpn*I (5′-ggtacc atgatgcagcagccaccaccc-3′) and *Hin*dIII (5′-aagctt ttagtagctaaacccgacttgttgctg-3′) restriction sites at the 5′ ends (indicated by a underline in the primer sequence). PCR products were cloned into pGEM-T Easy Vector (Promega) and transformed into DH5 alpha cells. Inserts were sequenced to ensure that there were no errors during PCR amplification. The insert was subcloned by digesting with *Kpn*I and *Hin*dIII and ligating into the pET-30a (+) expression vector. The plasmid pET30a+AtRBP45b was transformed into BL21 cells. Protein expression was induced by the addition of 0.1 mM IPTG followed by incubation at 37°C for 3-4 h. Total protein extract from these cells was passed through the His-Select HF Nickel Affinity Gel (Sigma). The His-tagged recombinant protein bound to the column was eluted using the elution buffer (20 mM Tris pH 8.0, 500 mM NaCl, 250 mM imidazole and 10% glycerol) and was subjected to dialysis to remove the salts and glycerol. The splice variant form of AtRBP45b was amplified using gene specific primers that had *Bam*HI (5′-ggatcc atgatgcagcagccaccaccc-gga-3′) and *Xho*I (5′-ctcgag ttacttctcggaaaactgaacgaa-3′) sites at their 5′ ends. Recombinant splice variant protein was obtained the same way as described above.

MALDI-TOF analysis

Purified proteins were cut from Coomassie stained gels and processed on a MALDI-TOF mass spectrometer as described earlier (Jambunathan and Mahalingam, 2006).

Electrophoretic mobility shift assay

A U-rich intron of the rubisco activase (RCA) gene was PCR amplified using the primers RCA-F (5′-ttgactctgcctaacatcaagg-3′) and RCA-R (5′-ataagtggaacctatgagacaaaaa-3′) (Lorkovic et al., 2000) from Arabidopsis Col-0 ecotype genomic DNA. The 100 bp product was cloned into pGEM-T Easy Vector (Promega). The purified plasmid was digested with *Sal*I. The digested product was gel purified and phenol:chloroform extracted. Approximately 2-3 μg of purified, digested plasmid was transcribed using the Ambion MEGAshortscript™ Kit. Two microliters of P^{32}-UTP was added to radiolabel the transcript. RNA binding assays were conducted with purified recombinant protein of full-length and splice variant AtRBP45b. The protocol for the binding assays were as described earlier (Black et al., 1998) with a few modifications: about 1 μl 10× binding buffer (500 mM Tris-HCl pH 8.0, 1 M NaCl, 144 mM 2-mercaptoethanol), 1 μl radiolabeled *RCA* RNA (approximately 4 fM), and 3 μl nuclease free water were incubated at 95°C for three minutes and then placed on ice to denature the RNA. On ice, 1-5 μl of purified recombinant protein (~4-30 nM) was combined with MTPBS (150 mM NaCl, 16 mM Na_2HPO_4, 4 mM NaH_2PO_4 pH 7.3) to a total volume of 5 μl. RNA and protein solutions were combined and incubated at room temperature for 10 min. Samples were then UV crosslinked for 10 min using a Stratalinker. After the addition of 5 μl of loading dye, the samples were electrophoresed on a 4% polyacrylamide (80:1 acrylamide: bis acrylamide) gel in 0.5X TBE at 200 V at 4°C until the dye just ran off the gel. The gel was dried and analyzed using a PhosphoImager.

RESULTS

Plant RBPs with triple RRMs

To explore the RBPs with triple RRMs in plants, we first identified these genes in the well-curated Arabidopsis genome (http://www.arabidopsis.org). In a previous study, a list of 196 RRM-containing Arabidopsis proteins that contained both RNP1 and RNP2 submotifs was reported (Lorkovic and Barta, 2002). In this set, 24 proteins were identified as having three

RRM domains, which became the starting point of our analysis. A search for "RRM" domain-containing proteins in the Plant-RBP database (http://plantrbp.uoregon.edu) identified 304 putative orthologous groups (POGs) (Walker et al., 2006). A systematic analysis through each of these POGs identified 39 Arabidopsis proteins with multiple RRM domains. Based on both the earlier study (Lorkovic and Barta, 2002) and our POG database analysis, 24 Arabidopsis and 19 rice proteins were identified as having triple RRMs and were selected for further analysis (Table 1). For Arabidopsis and rice, the loci that consistently clustered together in both the POG and Genome cluster database (GCD; http://bioweb.ucr.edu/databaseWeb/index.jsp) (Horan et al., 2005) were selected. Poplar sequences with 3 RRMs were retrieved using the *Populus trichocarpa* v 1.1 page (http://genome.jgi-psf.org/Poptr1_1/Poptr1_1.home.html). Four hundred and five hits were obtained with "RNA + Recognition + Motif" as a search query under the advanced search option. The sequence of each of these proteins was examined individually for the presence of three RRMs. The Populus database shows conserved domains as predicted by HMMPfam, HMMSmart, and ProfileScan. If there were discrepancies between any of the prediction algorithms, the sequences were input into the NCBI's conserved domain database.

The 80 RBP sequences were subjected to a phylogenetic analysis. In the phylogenetic tree (Fig. 1), there were nine clades with four or more sequences. Seven of the clades contained representatives from all the three plant species. The largest clade consisted of 23 sequences, accounting for nearly 30% of the RBPs with triple RRMs. All the rice RBP sequences had a homolog in Arabidopsis or Poplar, while two sequences in Arabidopsis (At1g45100 and At1g34140) and two in Poplar (Poptr251120 and Poptr549854) did not have a closely related homolog. RBPs in clades 8 and 9 did not have representatives from the rice genome, suggesting that these sequences may be unique to eudicots.

There was a bias in the chromosomal distribution of Arabidopsis RBPs, with 12 of them located on chromosome 1. Of the eight Arabidopsis sequences in the largest clade, four of them were located on chromosome 1, with two of them positioned next to one another. Such chromosomal biases were not observed among the rice and poplar RBPs.

Comparison of conserved motifs in plant RBPs with triple RRMs

To further identify other motifs in these protein sequences, Multiple Em for Motif Elicitation (MEME) was used (Bailey et al., 2006) (Fig. 2). Motifs 1 (IFVGNLDKNVTEEDLRQLFEQYGEIV HCK), 2 (TGRSKGYGFVRFRDENEAQRA), and 5 (KGCGFVQ YYNRRCAEYAIQKL) were identified in all 80 sequences. Motif 3 (EVKTIWVGNLPHWMDENYLHSCFHHYGEI) was identified in 73 sequences, while motif 4 (VSIKVIRNKQTGQSEGYGFI EFYSHAAAEKVLQTYNGTMMPN) was present in 24 proteins. The underlined portions of these motifs show homology to either the RNP1 or RNP2 consensus sequences.

Further analysis of these sequences using the NCBI's conserved domain search database (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi) identified motifs 1 (P-value: 1e-06), 2 (P-value: 0.004), 3 (P-value: 0.006) and 4 (P-value: 5e-06) as RRMs, while motif 5 did not find any significant homologies to any other domains in the database. Pair-wise comparisons of these motifs showed a significant correlation between motifs 1 and 3 (Supplementary Table S2.)

The majority of the sequences belonging to a clade showed the same pattern or arrangement of these conserved motifs (Fig. 2). This prompted the Motif Alignment and Search Tool (MAST) analysis to identify proteins in the PubMed database that have similar arrangement of these five motifs. Proteins with significant matches to the various conserved motifs present in plant RBPs are shown in Table 2. Not surprisingly, all the proteins identified by this analysis were RBPs involved in various aspects of RNA metabolism during development and stress signaling.

Table 1. Locus identifiers of the plant RNA binding proteins with triple RNA Recognition Motifs

Arabidopsis thaliana	Oryza sativa	Populus trichocarpa
At1g03457	Os01g71200	Poptr714870
At1g11650	Os01g71770	Poptr554690
At1g17370	Os02g21430	Poptr410877
At1g34140	Os02g35950	Poptr645804
At1g43190	Os02g48790	Poptr203044
At1g45100	Os03g07800	Poptr583200
At1g47490	Os03g37270	Poptr673211
At1g47500	Os04g37690	Poptr663393
At1g49600	Os04g45930	Poptr286862
At1g54080	Os07g33330	Poptr651076
At1g60900	Os07g42380	Poptr281598
At2g29400	Os08g09100	Poptr292764
At2g43410	Os08g40880	Poptr256575
At2g44710	Os09g28810	Poptr837528
At3g14100	Os10g06130	Poptr199846
At3g19130	Os11g04390	Poptr293793
At3g52660	Os11g14430	Poptr767539
At4g00830	Os11g40510	Poptr713522
At4g03110	Os12g04180	Poptr549854
At4g27000		Poptr219589
At4g36960		Poptr744254
At5g09880		Poptr741275
At5g19350		Poptr286468
At5g54900		Poptr412314
		Poptr832343
		Poptr251120
		Poptr245147
		Poptr798901
		Poptr727110
		Poptr802569
		Poptr246020
		Poptr799273
		Poptr717692
		Poptr716366
		Poptr410596
		Poptr424261
		Poptr719459

Expression analysis of Arabidopsis RBPs with triple RRMs in response to ozone-induced oxidative stress

We focused our experiments on the largest clade, which contained eight Arabidopsis RBPs with triple RRMs. This group of

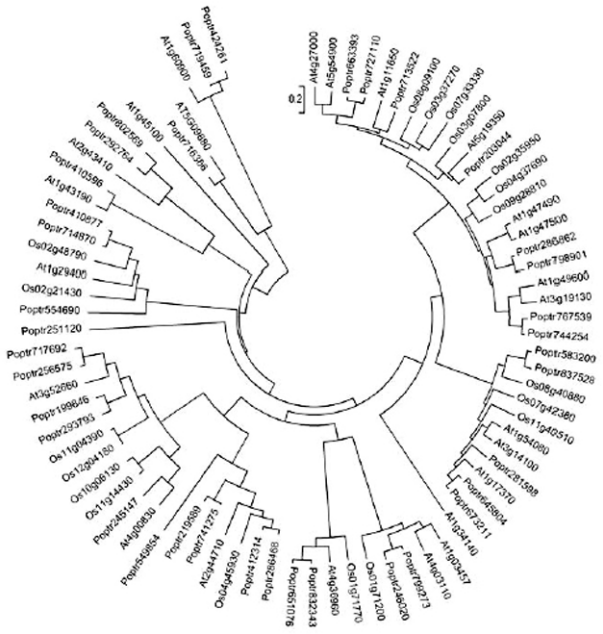

Fig. 1. Unrooted phylogenetic tree of 80 plant RNA binding proteins with triple RNA recognition motifs. The entire protein sequences of the RBPs were aligned by ClustalXI.83. The phylogenetic tree was constructed using the neighbor joining method in MEGA 4.0.

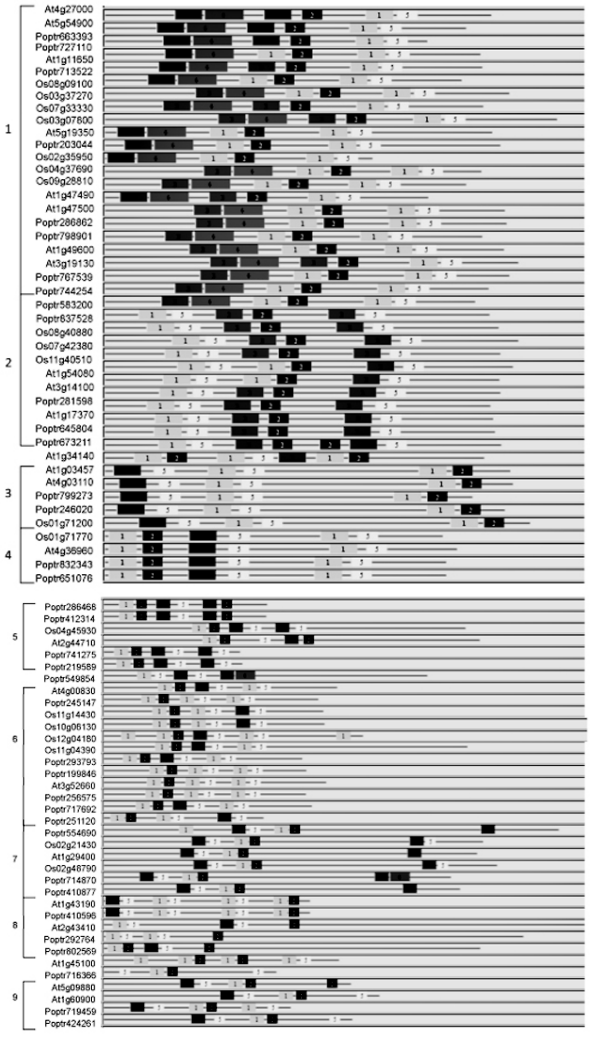

Fig. 2. Motif organization of the plant RNA binding proteins using Multiple Em for motif elicitation program. The groupings on the left correspond to the clade numbers based on the phylogenetic tree.

proteins is similar to the tobacco RBP45 and RBP47 proteins (Lorkovic et al., 2000), which are up-regulated in response to hydrogen peroxide in Arabidopsis suspension cultures (Schmidt et al., 2009). Furthermore, the protein encoded by At1g11650, AtRBP45b, showed approximately 60% homology to the yeast triple RRM containing RBP, CSX1, a global regulator of oxidative stress (Supplementary Fig. S1) (Rodriguez-Gabriel et al., 2003). CSX1 also showed the same arrangement of the five motifs (4-3-2-1-5) observed among the plant RBPs in the largest clade. Prompted by these sequence similarities to proteins involved in oxidative signaling, we examined the expression of these eight AtRBPs in response to acute ozone exposure (Fig. 3).

The steady-state levels of *At1g11650* (*AtRBP45b*) showed an initial decline at 1 h and an increased expression at 8, 16, and 24 h in response to ozone fumigation. The transcript levels of *At5g54900 (AtRBP45a)* increased significantly at the later time points. Steady-state levels of *At5g19350* were low compared to the other *AtRBP45*-related genes and did not show any significant changes in expression in response to ozone. Similarly, the transcript levels of *At4g27000 (AtRBP45c)* did not show any significant changes in response to ozone treatment.

Following ozone stress, *AtRBP47a (At1g49600)* showed an opposite pattern of expression when compared with members of the RBP45 group (Fig. 3). Transcript levels of this gene showed a sharp decrease at 12, 16 and 24 h. Steady state levels of *At3g19130 (AtRBP47b)* showed an increase at 1 but were strongly repressed at other time points. Transcript levels of *At1g47500 (AtRBP47c')* did not show any significant changes in response to ozone. Expression of *At1g47490 (AtRBP47c)* was hardly detectable under the hybridization and washing conditions used in this study (data not shown).

AtRBP45b is constitutively expressed and is alternatively spliced

Through Northern blot analysis, *AtRBP45b* was found to be expressed in roots, leaves, stems, flowers and siliques. The highest relative expression was observed in the leaves. The lowest relative expression was observed in the root tissues (Fig. 4). This is consistent with the data retrieved from the Genvestigator (Supplementary Fig. S2).

Gene models for *AtRBP45b (At1g11650.2)* indicated that there are six exons and five introns, yielding a protein with 405 amino acids (Fig. 5). Another gene model (*At1g11650.1*) is predicted to be alternatively spliced at the proximal 3′ splice site of intron 4, yielding a protein with 305 amino acids due to the introduction of a premature stop codon. A 3′ RACE PCR was conducted to verify these informatics-based predictions. Similar to the prediction, a splice variant was found that was truncated at the fourth exon, which resulted in a premature stop codon. Thus, the splice variant had only a portion of the third RRM motif. However, the two different splice variants identified had differences in their 3′ UTR sequences that were not predicted in the database (Fig. 5). The *AtRBP45b-SV1* retained intron 4, while *AtRBP45b-SV2* retained both intron 4 and intron 5.

Table 2. Summary of motif alignment and search tool analysis of plant RBPs with triple RRMs

Clade[a]	Conserved motifs[b]	Homologous protein[c]	E-value
1	3-4-1/3-2-1-5	CSX1 related protein (4-1-2-5) Early lethal abnormal vision (ELAV) protein (4-1-2-1-5)	1.6e-27
2	1-5-1/3-2-3-5	T-cell induced antigen (TIA) and and TIA1- related (TIAR) (1-5-1-2-1-5)	2.1e-33
3	3-5-1-5-1-2	BRUNO (1-5-1-5-1-2)	5.4e-28
4	1-2-3-5-1-5	Nucleolin (1-2-15-1-5)	7.2e-09
5	1-2-3-5-3-2/5	PolyA binding protein (PABP) (1-2-1-5-1-2)	4.3e-25
6	1-2-1/3-5-1-5	Heterogeneous ribonucleoprotein Q (hnRNP Q) (1-2-3-5-1-5)	6.7e-25
7	3-5-1-2	Meiosis-related protein 2 (Mei2) (1-5-1-2)	1.3e-24
8	3-5-1-5-1-2-1-5	RNA binding motif protein 39 (Rbm39) (1-5-1-2-5)	1.4e-30
9	3-5-1-2-2/5	Splicing factor (1-5-1-2-1-5)	7.1e-23

[a]The clade numbers are shown in Fig. 2.

[b]Organization of the five motifs identified by MEME analysis in majority of RBPs of each clade.

[c]The best protein match identified by MAST analysis. The organization of the motifs when different from the ones identified in plant RBPs is shown in paranthesis.

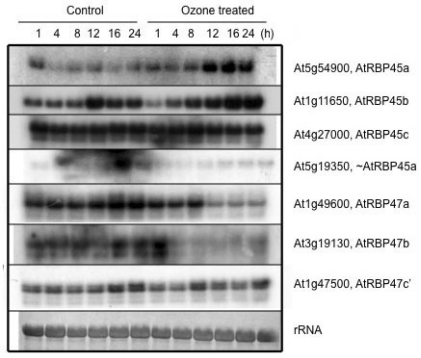

Fig. 3. Northern blot hybridizations using the gene specific products of the Arabidopsis RNA binding proteins belonging to the RBP45/47 subgroup. The locus identifiers of the genes are shown on the right. A picture of the methylene blue staining of rRNA is shown in the bottom panel.

Full-length and splice variant forms of AtRBP45b protein bind RNA

Full-length and splice variant proteins of AtRBP45b were generated as His-tagged fusions using the pET vector system. Following purification on nickel affinity columns, single discrete protein bands of approximately 54 kDa and 44 kDa were observed for the full-length and the splice variant forms, respectively (Supplementary Fig. S4). Taking into account the 10 kDa for the His-tag, the observed size of the proteins matched their predicted sizes. The fusion proteins were subjected to MALDI-TOF analysis that showed 15 out of 28 matches (33% coverage, Score: 146, Expect 8.1e-09) to the predicted full-length AtRBP45b, and 10 out of 20 matches (30% coverage, Score: 122, Expect 2e-06) to the predicted splice variant, confirming the purity of these recombinant proteins.

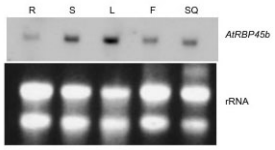

Fig. 4. RNA gel blot showing the expression of *AtRBP45b* gene in various plant tissues. A picture of the methylene blue staining of rRNA is shown in the bottom panel. R-roots, S-stem, L-leaves, F-flowers, SQ-siliques.

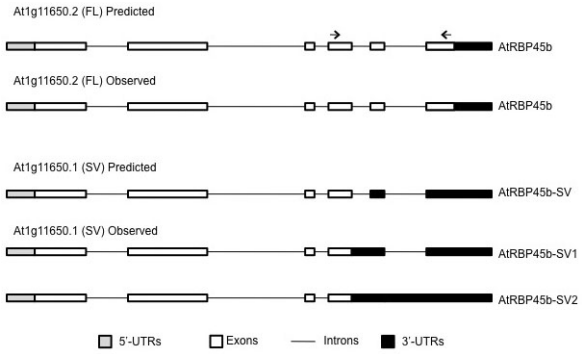

Fig. 5. Gene models of *AtRBP45b.* Observed gene models were based on the results of 3′ RACE PCR analysis. Primer pairs used for RT- PCR analysis are indicated by arrows.

The U-rich sequence of the *rca1* gene intron was used to demonstrate the RNA binding property of the two purified recombinant proteins by electrophoretic mobility shift assay (Fig. 6). This particular sequence was previously used to demonstrate the RNA binding property of the tobacco RBP45/47 proteins (Lorkovic et al., 2000). Both full-length as well as the splice variant forms were able to bind to this RNA. For both proteins, binding was undetectable at protein concentrations below 10 nanomolar and almost completely saturating at concentrations above 12 nanomolar. Since the splice variant form bound to the *rca1*RNA, the third RRM domain may not be crucial for RNA binding.

Full-length and splice variant forms of *AtRBP45b* are differentially expressed in response to biotic and abiotic stresses

We examined the expression of the full-length and the two splice variants of *AtRBP45b* in response to direct infiltration of oxidants, abiotic stresses, and biotic stresses at multiple time points (Fig. 7). The full-length *AtRBP45b* and the splice variant form *AtRBP45b-SV2* were constitutively expressed. Interestingly, the AtRBP45b-SV1 isoform was observed only in response to stressors that were caused by infiltration of chemicals or pathogens, while salt and temperature stresses that were imparted without any mechanical perturbations failed to do so. This indicated that *AtRBP45b-SV1* is rendered stable in response to wounding.

The relative abundance of the splice variants was approximately less than 10% of the full-length transcript (Fig. 7). In order to determine the differential expression of the full-length

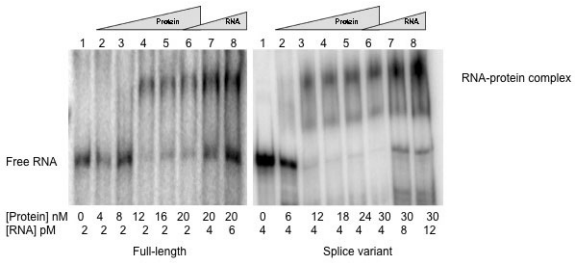

Fig. 6. Gel electrophoretic mobility shift assays demonstrating RNA-binding activity of recombinant AtRBP45b and AtRBP45b-SV. Concentrations of the protein and labeled *rca1* RNA used for each lane is shown below.

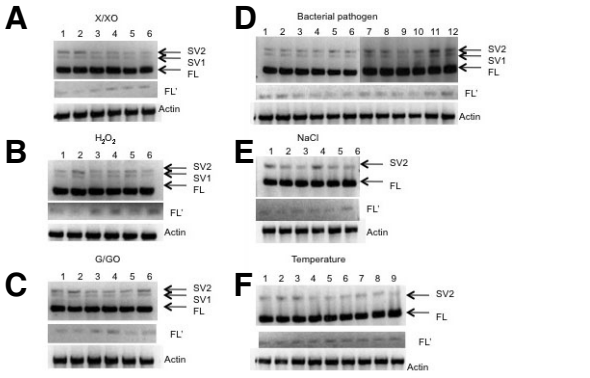

Fig. 7. RT-PCR analysis of *AtRBP45b* full-length gene and the two splice variants in response to abiotic and biotic stresses. Top panel shows amplification products with 3 μl of cDNA and 30 cycles of PCR. Arrows indicate AtRBP45b full length (FL) and the two splice variant forms (SV1 and SV2). Middle panels labeled as FL' are amplification products of full-length AtRBP45b with 1 μl of cDNA and 20 cycles of PCR. The bottom panel represents the amplification of constitutively expressed actin gene showing the quality of the cDNA. (A) Xanthine/Xanthine oxidase (X/XO) treatment. Lane 1, 1 h X, lane 2, 1 h X/XO; lane 3, 8 h X; lane 4, 8 h X/XO; lane 5, 24 h X; lane 6, 24 h X/XO. (B) Hydrogen peroxide treatment. Lane 1, 1 h water; lane 2, 1 h H_2O_2; lane 3, 8 h water; lane 4, 8 h H_2O_2. Lane 5, 24 h water; lane 6, 24 h H_2O_2. (C) Glucose/Glucose oxidase (G/GO) treatment. Lane 1, 1 h G; lane 2, 1 h G/GO; lane 3, 8 h G; lane 4, 8 h G/GO; lane 5, 24 h G; lane 6, 24 h G/GO. (D) Bacterial pathogen treatment. Virulent strain *P. syringae* DC3000 and avirulent strain *P. syringae* DC3000 + AvrRpt2. Lane 1, 1 h $MgCl_2$; lane 2, 1 h DC3000; lane 3, 1h AvrRpt2; lane 4, 2 h $MgCl_2$; lane 5, 2 h DC3000; lane 6, 2 h AvrRpt2; lane 7, 4 h $MgCl_2$; lane 8, 4 h DC3000; lane 9, 4 h AvrRpt2; lane 10, 6 h $MgCl_2$; lane 11, 6 h DC3000; lane 12, 6 h AvrRpt2. (E) Salt stress treatment. Lane 1, 1 h water; lane 2, 1 h NaCl; lane 3, 8 h water; lane 4, 8 h NaCl; lane 5, 24 h water; lane 6, 24 h NaCl. (F) Temperature stress treatment. lane 1, 1 h Control; lane 2, 1 h Heat; lane 3, 1 h Cold; lane 4, 8 h Control; lane 5, 8 h Heat; lane 6, 8 h Cold; lane 7, 24 h Control; lane 8, 24 h Heat; lane 9, 24 h Cold.

transcript, RT-PCR was restricted to 20 cycles (FL'), as opposed to the 30 cycles used for the detection of the two splice variants. The differences in the abundances of the full-length transcript observed in response to the various treatments were not significant.

DISCUSSION

The RNA recognition motif (RRM) is the most common domain found exclusively in eukaryotic proteins involved in RNA metabolism (Anantharaman et al., 2002). In plants, the superfamily of proteins containing the RRM has vastly expanded compared to other multi-cellular organisms, such as Drosophila and *C. elegans*, and, in fact, is comparable in size to that of humans (Anantharaman et al., 2002). In a previous study, it was reported that there are 196 RRM-containing proteins in Arabidopsis (Lorkovic and Barta, 2002). Based on this study using the POG database, 261 RBPs containing one or more RRMs were identified. Lorkovic and Barta (2002) used protein sequences of known plant or metazoan proteins to query the Arabidopsis database using the BLAST algorithm. Our study employed mutual-best-hit and phylogenetic approaches in the POG database, to identify orthologous sequences (Walker et al., 2006). We surmise that this difference in approach between the two analyses may account for the differences in the number of proteins identified.

Of the 261 RBPs, 157 proteins contained a single RRM, 70 had two RRMs, 24 had three RRMs, eight had four RRMs and two proteins had five RRM domains. The 24 RBPs with three RRMs identified in the earlier report (Lorkovic and Barta, 2002) and our analysis using the POG/RBP database were identical. RBPs with triple RRMs have been implicated in various facets of RNA metabolism, including RNA stability and splicing, and show similarities to some metazoan hnRNPs (Lorkovic and Barta, 2002). These proteins were also shown to be involved in flower development (Park et al., 2006; Schomburg et al., 2001).

Identification of orthologs of Arabidopsis RBPs with three RRMs in Poplar (a perennial) and rice showed not only that these proteins are conserved in plants but also that they may play important roles in normal plant growth and development (Table 1). Phylogenetic analysis of these 80 RBP sequences showed that only two sequences from Arabidopsis and two sequences from Poplar were singletons, suggesting that RBPs with triple RRMs have been well-conserved during plant evolution. Consistent with the expansion of all gene families in the Populus genome (Tuskan et al., 2006), the number of RBPs with triple RRMs was nearly twice as many as in the rice genome. There were 10 RBPs identified from the Arabidopsis and/or Poplar genomes (belonging to clades 8 and 9) that did not have corresponding rice orthologs. Further analysis is warranted to determine if these proteins are truly unique to the eudicots.

It has been well documented that large segments of plant genomes consist of duplicate loci due to high frequencies of tandem duplications (Jain et al., 2006), segmental duplications (Kim et al., 2006), and polyploidization events (Adams and Wendel, 2005; Moore and Purugganan, 2005). The clustering of Arabidopsis RBPs on Chromosome 1 could be due to tandem and/or segmental duplications. The dispersed occurrence of RBPs in rice and poplar supports the notion of ancient polyploidization events in these genomes (Paterson et al., 2004; Sterck et al., 2005).

Analysis of the protein sequences using MEME showed that motifs 1, 2, 3 and 4 shared significant homologies to the RRM signature and were present in various combinations in these proteins, further confirming that all 80 sequences had three RRMs. Interestingly, the motif 5 sequence, which was identified in all 80 proteins, did not show homology to any sequence in the conserved domain database. Careful examination of the motif 5 sequence did show the presence of the RNP1 signature, suggesting that our search parameters for motif identification in

MEME did not identify any novel motifs in the plant RBPs with triple RRMs.

Further sequence analysis using MEME (Fig. 2) and MAST programs aided in associating the plant RBPs with other eukaryotic RBPs based on the organization of the five conserved motifs (Table 2). The largest clade, containing 23 RBPs, was highly homologous to both the CSX1 protein in fission yeast, an RBP with a triple RRM and described as a global regulator of oxidative stress (Rodriguez-Gabriel et al., 2003) and the ELAV family of proteins, which are important for cell differentiation and alternate splicing (Antic and Keene, 1997). The AtRBP45b (At1g11650) protein sequence showed high homology (59% similarity) to the yeast CSX1 protein (Supplementary Fig. S1).

Other clades of plant RBPs with triple RRMs showed strong homologies to proteins involved in various facets of RNA metabolism. Plant RBPs in the second clade showed strong homologies to T-cell-restricted intracellular antigen (TIA-1) and TIA-1-related protein (TIAR), both of which function as mRNA turnover and translation regulatory (TTR) RBPs (Mazan-Mamczarz et al., 2006; Pullmann et al., 2007). TIA-1 and TIAR also are involved in signaling apoptotic cell death in mammalian systems (Kawakami et al., 1992; Taupin et al., 1995; Tian et al., 1991) and are important players in oxidative stress signaling (Abdelmohsen et al., 2008).

Clade 3 proteins showed significant homology to Bruno-like RBPs, which are important for binding CUG trinucelotide repeats and are implicated in human diseases and differentiation of cell types (Good et al., 2000). Bruno proteins are well conserved in multi-cellular organisms and act as translational repressors that play a key role during gametogenesis (Webster et al., 1997). Clade 4 RBPs had a motif organization that was also identified among nucleolins important for ribosome biogenesis (Srivastava and Schlessinger, 1990). Nucleolins are also cell surface receptors, autoantigens, and transcriptional repressors (Srivastava and Pollard, 1999). Clade 5 proteins showed high homologies to several poly-A binding proteins (PABPs) that are important for RNA stability and translation. As central regulators of gene expression, PABPs are targeted by a variety of DNA and RNA viruses, which leads to changes in complex formation, RNA stability, and localization (Smith and Gray, 2010). RBPs belonging to clades 6 and 9 showed motif organization similar to the hnRNP Q-like proteins that are important for mRNA splicing (Singh, 2007). Apart from splicing, hnRNPs are important for mRNA packaging, transport, and silencing (Singh, 2001). Some hnRNPs are also essential for gametogenesis (Kinnaird et al., 2004). Clade 7 proteins showed homology to the meiosis regulatory proteins (Kaur et al., 2006). RBPs in Clade 8 showed homology to proteins involved in mRNA splicing and also to those that act as transcriptional coactivators (Jung et al., 2002). Based on these similarities to well characterized eukaryotic RBPs, it is obvious that plant RBPs with triple RRMs have diverse cellular functions.This analysis provides a good starting point to examine the precise functional roles of each of these proteins in plants.

Using the Genevestigator program, digital northern analysis of the 24 Arabidopsis RBPs with triple RRMs further demonstrated that many of these genes are expressed in various tissues and in different stages of development (Supplementary Fig. S2), indicating that they play very important roles throughout the lifecycle of a plant. Many of the Arabidopsis RBPs were transcriptionally responsive to biotic, abiotic, chemical, and hormonal treatments, suggesting an involvement in stress signal transduction (Supplementary Fig. S3). Hierarchical clustering of RBP genes along with their gene expression profiles showed that RBPs from different clades grouped together (Supplementary Fig. S3). This suggested that, following gene duplications, one of the duplicates underwent mutations leading to pseudogenization or neofunctionilization (Lynch and Conery, 2000).

Northern analysis further confirmed the gene expression changes in response to ozone-induced oxidative stress among the eight Arabidopsis RBPs that are similar in sequence to the yeast CSX1 (Fig. 3). Interestingly, two members of the RBP45 group and two members of RBP47 group show opposite patterns of expression. In mammalian systems, some of the TTR-RBPs with contrasting gene expression patterns have opposing post-transcriptional influence on the same set of target genes (Abdelmohsen et al., 2008). It is tempting to speculate that such a phenomenon may be operative in plants.

Among the members of the AtRBP45 group, we focused our attention on the *AtRBP45b* gene since it was constitutively expressed in several different tissues (Fig. 4), consistent with earlier reports (Lorkovic et al., 2000; Park et al., 2006). Since this gene also was induced in response to oxidative stress, we speculate it may be crucial for coordinating the adaptive response of plants to stresses in different organs and different developmental stages. The 3′ RACE analysis enabled the identification of two new splice variant forms (designated as *AtRBP45b-SV1* and *AtRBP45b-SV2*) of the *AtRBP45b* gene (Fig. 5). The predicted splice variant form is due to an alternative 3′ splice site that causes the first three nucleotides, "taa," of the fourth intron to be included, thus resulting in a truncated protein lacking 100 amino acids at the C-terminus. Both the variants were different when compared with the predicted splice variant for this gene found in the Arabidopsis database. Our data suggested that the retention of the entire fourth intron resulted in a truncated protein due to the premature stop codon. The difference between the two new splice variants was due to the last intron, which was retained in *AtRBP45b-SV2* but was spliced out in *AtRBP45b-SV1* (Fig. 5). It is not known whether these splice variant forms are translated *in planta*. In higher organisms, many transcripts with a termination codon located more than 50 nucleotides upstream of an exon-exon junction are targets of nonsense-mediated decay (NMD) (Green et al., 2003; Lewis et al., 2003). In plants, the splice variants of the SR family of RBPs meet these criteria and are speculated to be ideal targets for the NMD pathway (Palusa et al., 2007). Similarly, the splice variants of the *AtRBP45b* meet these criteria and, hence, may be degraded selectively by NMD.

The binding of the U-rich *rca1* intron by both the full length as well as the splice variant of recombinant AtRBP45b (Supplementary Fig. S4; Fig. 6) is consistent with a previous report on the tobacco RBP45 protein (Lorkovic et al., 2000). Deletion derivatives of the tobacco RBP45 protein were able to bind to the *rca1* RNA when any combination of two RNA binding domains was present (Lorkovic et al., 2000). Along the same lines, we infer that the AtRBP45b splice variant is able to bind to RNA with two full RRMs and that the truncated third RRM may not affect binding (Fig. 6).

Based on the strong similarities in sequence and constitutive expression patterns we surmise that AtRBP45b is a functional homolog of the tobacco RBP45. Similar to the tobacco RBP45 protein, the Arabidopsis AtRBP45b was shown to be associated with poly-A RNA (Lorkovic et al., 2000; Schmidt et al., 2009). Protein interaction database analysis indicated that the AtRBP45b interacts with the cap binding protein CBP20 (Supplementary Table S3). We speculate that AtRBP45b most likely binds to the U-rich sequences with the 3'-UTRs of target genes and, in co-ordination with the CBP20, may be involved in regulating gene expression during normal development and in re-

sponse to environmental perturbations.

RT-PCR analysis demonstrated that the splice variant *AtRBP45b-SV2* was constitutively expressed while the *AtRBP45b-SV1* was observed only during mechanical perturbations. Since the alternate splice variants are presumably produced through transcription of the full-length RNA transcript via a single promoter, we speculate that *AtRBP45b-SV1* transcripts are rendered stable during mechanical perturbations. Recently, members of the UBA2 family of RBPs have been shown to be induced in response to wounding (Bove et al., 2008; Kim et al., 2008), supporting a role for the RBPs in senescence and stress signaling pathways. Understanding the precise function of the AtRBP45b protein and its splice variant forms during stress signaling entails the identification of the target RNA molecules that they bind under different conditions through the use of techniques like ribonucleoprotein immunoprecipitation assays in conjunction with DNA microarrays (Niranjanakumari et al., 2002; Schmitz-Linneweber et al., 2005). These assays, coupled with over-expressor and/or knockout mutants of *AtRBP45b*, will shed light on post-transcriptional gene regulatory processes operative during both normal development and stress environments.

Note: Supplementary information is available on the Molecules and Cells website (www.molcells.org).

ACKNOWLEDGMENTS

This research was supported in part by funds from Oklahoma Agricultural Experiment Station. L.P. is supported by the Oklahoma Lois Stokes Alliance for Minority Participation Bridge to Doctorate program NSF-HRD-0402640.

REFERENCES

Abdelmohsen, K., Kuwano, Y., Kim, H.H., and Gorospe, M. (2008). Posttranscriptional gene regulation by RNA-binding proteins during oxidative stress: implications for cellular senescence. Biol. Chem. *389*, 243-255.

Adams, K.L., and Wendel, J.F. (2005). Polyploidy and genome evolution in plants. Curr. Opin. Plant Biol. *8*,135-141.

Adams, M.D., Celniker, S.E., Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Scherer, S.E., Li, P.W., Hoskins, R.A., et al. (2000). The genome sequence of Drosophila melanogaster. Science *287*, 2185-2195.

Anantharaman, V., Koonin, E.V., and Aravind, L. (2002). Comparative genomics and evolution of proteins involved in RNA metabolism. Nucleic Acids Res. *30*, 1427-1464.

Antic, D., and Keene, J.D. (1997). Embryonic lethal abnormal visual RNA-binding proteins involved in growth, differentiation, and posttranscriptional gene expression. Am. J. Hum. Genet. *61*, 273-278.

Bailey, T.L., and Gribskov, M. (1998). Combining evidence using p-values, application to sequence homology searches. Bioinformatics *14*, 48-54.

Bailey, T.L., Williams, N., Misleh, C., and Li, W.W. (2006). MEME: discovering and analyzing DNA and protein sequence motifs. Nucleic Acids Res. *34*, W369-W373.

Barkan, A., Klipcan, L., Ostersetzer, O., Kawamura, T., Asakura, Y., and Watkins, K.P. (2007). The CRM domain: An RNA binding module derived from an ancient ribosome-associated protein. RNA *13*, 55-64.

Black, D.L., Cahn, R., Min, H., Wang, J., and Bell, L. (1998). The electrophoretic mobility shift assay for RNA binding proteins. In RNA: Protein Interactions: A Practical Approach, C.W.J. Smith, ed. (New York Tokyo: Oxford University Press), pp. 109-135.

Bove, J., Kim, C.Y., Gibson, C.A., and Assmann, S.M. (2008). Characterization of wound-responsive RNA-binding proteins and their splice variants in Arabidopsis. Plant Mol. Biol. *67*, 71-88.

Burd, C.G., and Dreyfuss, G. (1994). Conserved structures and diversity of functions of RNA-binding proteins. Science *265*, 615-621.

Consortium, C.e.S. (1998). Genome sequence of the nematode C-elegans: A platform for investigating biology. Science *282*, 2012-2018.

Danon, A., and Mayfield, S.P.Y. (1991). Light regulated translational activators - Identification of chloroplast gene specific messenger-RNA binding proteins. EMBO *10*, 3993-4001.

Fedoroff, N.V. (2002). RNA-binding proteins in plants: the tip of an iceberg? Curr. Opin. Plant Biol. *5*, 452-459.

Fu, Z.Q., Guo, M., Jeong, B.R., Tian, F., Elthon, T.E., Cerny, R.L., Staiger, D., and Alfano, J.R. (2007). A type III effector ADP-ribosylates RNA-binding proteins and quells plant immunity. Nature *447*, 284-288.

Gendra, E., Moreno, A., Alba, M.M., and Pages, M. (2004). Interaction of the plant glycine-rich RNA-binding protein MA16 with a novel nucleolar DEAD box RNA helicase protein from Zea mays. Plant J. *38*, 875-886.

Good, P.J., Chen, Q., Warner, S.J., and Herring, D.C. (2000). A family of human RNA-binding proteins related to the Drosophila Bruno translational regulator. J. Biol. Chem. *275*, 28583-28592.

Green, R.E., Lewis, B.P., Hillman, R.T., Blanchette, M., Lareau, L.F., Garnett, A.T., Rio, D.C., and Brenner, S.E. (2003). Widespread predicted nonsense-mediated mRNA decay of alternatively-spliced transcripts of human normal and disease genes. Bioinformatics *19*, 118-121.

Horan, K., Lauricha, J., Bailey-Serre,s J., Raikhel, N., and Girke, T. (2005). Genome cluster database. A sequence family analysis platform for Arabidopsis and rice. Plant Physiol. *138*, 47-54.

Hugouvieux, V., Kwak, J.M., and Schroeder, J.I. (2001). An mRNA cap binding protein, ABH1, modulates early abscisic acid signal transduction in Arabidopsis. Cell *106*, 477-487.

Jain, M., Tyagi, A.K., and Khurana, J.P. (2006). Genome-wide analysis, evolutionary expansion, and expression of early auxin-responsive SAUR gene family in rice (*Oryza sativa*). Genomics *88*, 360-371.

Jambunathan, N., and Mahalingam, R. (2006). Analysis of arabidopsis growth factor gene 1 (GFG1) encoding a nudix hydrolase during oxidative signaling. Planta *224*,1-11.

Jensen, K.H., Herrin, D.L., Plumley, F.G., and Schmidt, G.W. (1986). Biogenesis of photosystem-II Complexes - Transcriptional, translational, and posttranslational regulation. J. Cell Biol. *103*, 1315-1325.

Jung, J.W., Ji, A.R., Lee, J., Kim, U.J., and Lee, S.T. (2002). Organization of the human PTK7 gene encoding a receptor protein tyrosine kinase-like molecule and alternative splicing of its mRNA. Biochim. Biophys. Acta *1579*, 153-163.

Kaur, J., Sebastian, J., and Siddiqi, I. (2006). The Arabidopsis-mei2-like genes play a role in meiosis and vegetative growth in Arabidopsis. Plant Cell *18*, 545-559.

Kawakami, A., Tian, Q.S., Duan, X.C., Streuli, M., Schlossman, S.F., and Anderson P. (1992). Identification and functional characterization of a Tia-1-related nucleolysin. Proc. Natl. Acad. Sci. USA *89*, 8681-8685.

Kiledjian, M., and Dreyfuss, G. (1992). Primary structure and binding activity of the Hnrnp U-Protein - binding RNA through RGG box. EMBO *11*, 2655-2664.

Kim, J., Shiu, S.H., Thoma, S., Li, W.H., and Patterson, S.E. (2006). Patterns of expansion and expression divergence in the plant polygalacturonase gene family. Genome Biol. *7*, R87.

Kim, Y.O., Pan, S., Jung, C.H., and Kang, H. (2007). A zinc finger-containing glycine-rich RNA-binding protein, atRZ-1a, has a negative impact on seed germination and seedling growth of Arabidopsis thaliana under salt or drought stress conditions. Plant Cell Physiol. *48*, 1170-1181.

Kim, J.S., Jung, H.J., Lee, H.J., Kim, K.A., Goh, C.H., Woo, Y., Oh, S.H., Han, Y.S., and Kang, H. (2008). Glycine-rich RNA-binding protein 7 affects abiotic stress responses by regulating stomata opening and closing in *Arabidopsis thaliana*. Plant J. *55*, 455-466.

Kim, C.Y., Bove, J., and Assmann, S.M. (2008). Overexpression of wound-responsive RNA-binding proteins induces leaf senescence and hypersensitive-like cell death. New Phytol. *180*, 57-70.

Kim, J.Y., Kim, W.Y., Kwak, K.J., Oh, S.H., Han, Y.S., and Kang, H. (2009). Zinc finger-containing glycine-rich RNA-binding protein in Oryza sativa has an RNA chaperone activity under cold stress conditions. Plant Cell Environ. *33*, 759-768.

Kim, J.Y., Kim, W.Y., Kwak, K.J., Oh, S.H., Han, Y.S., and Kang, H.

(2010). Glycine-rich RNA-binding proteins are functionally conserved in Arabidopsis thaliana and Oryza sativa during cold adaptation process. J. Exp. Bot. *61*, 2317-2325.
Kinnaird, J.H., Maitland, K., Walker, G.A., Wheatley, I., Thompson, F.J., and Devaney, E. (2004). HRP-2, a heterogeneous nuclear ribonucleoprotein, is essential for embryogenesis and oogenesis in Caenorhabditis elegans. Exp. Cell Res. *298*, 418-430.
Kwon, C., and Chung, I.K. (2004). Interaction of an Arabidopsis RNA-binding protein with plant single-stranded telomeric DNA modulates telomerase activity. J. Biol. Chem. *279*, 12812-12818.
Lasko, P. (2000). The Drosophila melanogaster genome: Translation factors and RNA binding proteins. J. Cell Biol. *150*, 51-56.
Lewis, B.P., Green, R.E., and Brenner, S.E. (2003). Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. Proc. Natl. Acad. Sci. USA *100*, 189-192.
Lim, M.H., Kim, J., Kim, Y.S., Chung, K.S., Seo, Y.H., Lee, I., Kim, J., Hong, C.B., Kim, H.J., and Park, C.M. (2004). A new Arabidopsis gene, FLK, encodes an RNA binding protein with K homology motifs and regulates flowering time via FLOWERING LOCUS C. Plant Cell *16*, 731-740.
Lorkovic, Z.J. (2009). Role of plant RNA-binding proteins in development, stress response and genome organization. Trends Plant Sci. *14*, 229-236.
Lorkovic, Z.J., and Barta, A. (2002). Genome analysis: RNA recognition motif (RRM) and K homology (KH) domain RNA-binding proteins from the flowering plant *Arabidopsis thaliana*. Nucleic Acids Res. *30*, 623-635.
Lorkovic, Z.J., Kirk, D.A.W., Klahre, U., Hemmings-Mieszczak, M., and Filipowicz, W. (2000). RBP45 and RBP47, two oligouridylate-specific hnRNP-like proteins interacting with poly(A)(+) RNA in nuclei of plant cells. RNA *6*, 1610-1624.
Lu, C., and Fedoroff, N. (2000). A mutation in the Arabidopsis HYL1 gene encoding a dsRNA binding protein affects responses to abscisic acid, auxin and cytokinin. Plant Cell *12*, 2351-2366.
Lynch, M., and Conery, J.S. (2000). The evolutionary fate and consequences of duplicate genes. Science *290*, 1151-1155.
Macknight, R., Bancroft, I., Page, T., Lister, C., Schmidt, R., Love, K., Westphal, L., Murphy, G., Sherson, S., Cobbett, C., and Dean, C. (1997). FCA, a gene controlling flowering time in Arabidopsis, encodes a protein containing RNA-binding domains. Cell *89*, 737-745.
Mahalingam, R., Shah, N., Scrymgeour, A., and Fedoroff, N. (2005). Temporal evolution of the Arabidopsis oxidative stress response. Plant Mol. Biol. *57*, 709-730.
Mahalingam, R., Jambunathan, N., Gunjan, S.K., Faustin, E., Weng, H., and Ayoubi, P. (2006). Analysis of oxidative signalling induced by ozone in Arabidopsis thaliana. Plant Cell Environ. *29*, 1357-1371.
Manley, J.L., and Tacke, R. (1996). SR proteins and splicing control. Genes Dev. *10*, 1569-1579.
Marchler-Bauer, A., Anderson, J.B., Cherukuri, P.F., DeWweese-Scott, C., Geer, L.Y., Gwadz, M., He, S.Q., Hurwitz, D.I., Jackson, J.D., Ke, Z.X., et al. (2005). CDD: a conserved domain database for protein classification. Nucleic Acids Res. *33*, D192-D196.
Mazan-Mamczarz, K., Lal, A., Martindale, J.L., Kawai, T., and Gorospe, M. (2006). Translational repression by RNA-binding protein TIAR. Mol. Cell. Biol. *26*, 2716-2727.
Moore, R.C., and Purugganan, M.D. (2005). The evolutionary dynamics of plant duplicate genes. Curr. Opin. Plant Biol. *8*, 122-128.
Mussgnug, J.H., Wobbe, L., Elles, I., Claus, C., Hamilton, M., Fink, A., Kahmann, U., Kapazoglou, A., Mullineaux, C.W., Hippler, M., et al. (2005). NAB1 is an RNA binding protein involved in the light-regulated differential expression of the light-harvesting antenna of *Chlamydomonas reinhardtii*. Plant Cell *17*, 3409-3421.
Niranjanakumari, S., Lasda, E., Brazas, R., and Garcia-Blanco, M.A. (2002). Reversible cross-linking combined with immunoprecipitation to study RNA-protein interactions *in vivo*. Methods *26*, 182-190.
Palusa, S.G., Ali, G.S., and Reddy, A.S. (2007). Alternative splicing of pre-mRNAs of Arabidopsis serine/arginine-rich proteins, regulation by hormones and stresses. Plant J. *49*, 1091-1107.
Park, J.I., Endo, M., Kazama, T., Saito, H., Hakozaki, H., Takada, Y., Kawagishi-Kobayashi, M., and Watanabe, M. (2006). Molecular characterization of two anther-specific genes encoding putative RNA-binding proteins, AtRBPs, in *Arabidopsis thaliana*. Genes Genet. Syst. *81*, 355-359.
Paterson, A.H., Bowers, J.E., and Chapman, B.A. (2004). Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. Proc. Natl. Acad. Sci. USA *101*, 9903-9908.
Pullmann, R., Kim, H.H., Abdelmohsen, K., Lal, A., Martindale, J.L., Yang, X.L., and Gorospe, M. (2007). Analysis of turnover and translation regulatory RNA-Binding protein expression through binding to cognate mRNAs. Mol. Cell. Biol. *27*, 6265-6278.
Rochaix, J.D. (2001). Posttranscriptional control of chloroplast gene expression. From RNA to photosynthetic complex. Plant Physiol. *125*, 142-144.
Rodriguez-Gabriel, M.A., Burns, G., McDonald, W.H., Martin, V., Yates, J.R., Bahler, J., and Russell, P. (2003). RNA-binding protein Csx1 mediates global control of gene expression in response to oxidative stress. EMBO J. *22*, 6256-6266.
Saitou, N., and Nei, M. (1987). The neighbor-joining method - a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. *4*, 406-425.
Schmidt, F., Marnef, A., Cheung, M.K., Wilson, I., Hancock, J., Staiger, D., and Ladomery, M. (2009). A proteomic analysis of oligo(dT)-bound mRNP containing oxidative stress-induced *Arabidopsis thaliana* RNA-binding proteins ATGRP7 and ATGRP8. Mol. Biol. Rep. *37*, 839-845.
Schmitz-Linneweber, C., Williams-Carrier, R., and Barkan, A. (2005). RNA immunoprecipitation and microarray analysis show a chloroplast pentatricopeptide repeat protein to be associated with the 5′ region of mRNAs whose translation it activates. Plant Cell *17*, 2791-2804.
Schomburg, F.M., Patton, D.A., Meinke, D.W., and Amasino, R.M. (2001). FPA, a gene involved in floral induction in Arabidopsis, encodes a protein containing RNA-recognition motifs. Plant Cell *13*, 1427-1436.
Singh, O.P. (2001). Functional diversity of hnRNP proteins. Indian J. Biochem. Biophys. *38*, 129-134.
Singh, R.N. (2007). Evolving concepts on human SMN pre-mRNA splicing. RNA Biol. *4*, 7-10.
Smith, R.W., and Gray, N.K. (2010). Poly(A) binding protein (PABP): a common viral target. Biochem. J. *426*, 1-12.
Srivastava, A.K., and Schlessinger, D. (1990). Mechanism and regulation of bacterial ribosomal RNA processing. Annu. Rev. Microbiol. *44*, 105-129.
Srivastava, A.K., and Pollard, H.B. (1999). Molecular dissection of nucleolin's role in growth and cell proliferation: new insights. FASEB J. *13*, 1911-1922.
Sterck, L., Rombauts, S., Jansson, S., Sterky, F., Rouze, P., and Van de Peer, Y. (2005). EST data suggest that poplar is an ancient polyploid. New Phytol. *167*, 165-170.
Tamura, K., Dudley, J., Nei, M., and Kumar, S. (2007). MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. *24*, 1596-1599.
Taupin, J.L., Tian, Q.S., Kedersha, N., Robertson, M., and Anderson, P. (1995). The RNA-binding protein TIAR Is translocated from the nucleus to the cytoplasm during Fas-mediated apoptotic cell-death. Proc. Natl. Acad. Sci. USA *92*, 1629-1633.
Tian, Q.S., Streuli, M., Saito, H., Schlossman, S.F., and Anderson, P. (1991). A polyadenylate binding-protein localized to the granules of cytolytic lymphocytes induces DNA fragmentation in target cells. Cell *67*, 629-639.
Tuskan, G.A., Difazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., Putnam, N., Ralph, S., Rombauts, S., Salamov, A., et al. (2006). The genome sequence of black cottonwood, *Populus trichocarpa* (Torr. And Gray.). Science *313*, 1596-1604.
Walker, N.S., Stiffler, N., and Barkan, A. (2006). POGs/PlantRBP: aresource for comparative genomics in plants. Nucleic Acids Res. *35*, D852-D856.
Webster, P.J., Liang, L., Berg, C.A., Lasko, P., and Macdonald, P.M. (1997). Translational repressor bruno plays multiple roles in development and is widely conserved. Genes Dev. *11*, 2510-2521.
Zimmermann, P., Hennig, L., and Gruissem, W. (2005). Gene-expression analysis and network discovery using Genevestigator. Trends Plant Sci. *10*, 407-409.
Zuckerkandl, E., and Pauling, L. (1965). Evolutionary convergence and divergence in proteins. In Evolving Genes and Proteins, V. Bryson, and H.J. Vogel, eds. (New York, USA: Academic Press), pp. 97-166.