*Plant Molecular Biology* **30**: 1301–1306, 1996.



*Short communication*

# cDNA encoding a wheat (*Triticum aestivum* cv. Chinese Spring) glycine-rich RNA-binding protein

Mark J. Guiltinan* and Xiping Niu
*Department of Horticulture, The Biotechnology Institute, Intercollege Program in Plant Physiology, The Center for Gene Regulation, The Pennsylvania State University, 519 Wartik Laboratory, University Park, PA 16802, USA (*author for correspondence)*

Received 27 July 1995; accepted in revised form 8 December 1995

*Key words*: abscisic acid, glycine-rich, ribonucleoprotein, RNA-binding protein, RNP motif, wheat

**Abstract**

A wheat cDNA encoding a glycine-rich RNA-binding protein, whGRP-1, was isolated. WhGRP-1 contains two conserved domains, the RNA-binding motif (RNP motif) combined with a series of glycine-rich imperfect repeats, characteristic of a conserved family of plant RNA-binding proteins. Northern analysis revealed that whGRP-1 mRNA accumulates to high levels in roots and to lower levels in leaves of wheat seedlings. whGRP-1 mRNA accumulation is not enhanced by exogenous abscisic acid in seedlings and accumulates to very high levels during wheat embryo development, showing a pattern different from that of the ABA-inducible wheat Em gene.

RNA-binding proteins have been implicated in the control of gene expression at the levels of mRNA processing, transport, polyadenylation and RNA stability [10, 14]. As increasing numbers of RNA-binding proteins have been identified, several conserved RNA-binding motifs have been defined [5, 19]. The largest and best characterized group contains the RNP motif (sometimes called ribonucleoprotein consensus sequence (RNP-CS) or consensus sequence RNA-binding domain (CS-RBD)). The RNP motif consists of about 90 amino acids present in one or more copies per polypeptide. Two short highly conserved sequences, RNP1 (octamer) and RNP2 (hexamer) characterize this domain. In plants, RNP motif-containing RNA-binding proteins have been identified in spinach [28], tobacco [15–17, 20, 32], maize [7, 9, 11], *Arabidopsis thaliana* [2, 6, 23, 30], sorghum [8], carrot [29], mustard [13], and *Brassica napus* [4]. Within this group, a subfamily of ca. 16 kDa plant proteins consisting of a single RNP motif in combination with a glycine-rich imperfect repeat domain can be distinguished, represented by cDNAs isolated from maize, *Arabidopsis*, tobacco, sorghum, carrot, mustard, and *B. napus*. The plant glycine-rich RNA-binding proteins (GRPs, following the nomenclature first used [8]) have been

The nucleotide sequence data reported will appear in the GenBank, EMBL and DDBJ Nucleotide Sequence Databases under the accession number U32310 (whGRP-1).

implicated in many physiological processes, including ABA induction [11], water and chemical stress [9, 11], wounding [11, 29], and light and temperature-entrained circadian rhythm [6, 13]. Characterization of genes encoding RNA-binding proteins has revealed examples of developmental and tissue-specific regulation as well as tissue-specific alternative splicing of their products [6, 13, 15]. However to date, knowledge of the specific *in vivo* role of any of these proteins in plant development remains elusive.

We have isolated a full-length cDNA encoding a glycine rich nucleic acid-binding protein from wheat. Northern analysis revealed that the expression pattern of whGRP-1 is significantly different from that of the ABA-inducible Em gene [25], suggesting that unlike the maize RNA-binding protein gene MA16 [11], whGRP-1 expression is independent of ABA.

*Isolation of a cDNA encoding whGRP-1*

In the course of screening a wheat ABA-treated, stage III embryo cDNA library for clones encoding the DNA binding protein EmBP-1 [24], a phage was isolated that harbored a cDNA insert which upon detailed analysis, appeared to consist of two unrelated cDNAs artifactually fused head to head. At the 5′ end of this insert, upstream of an apparently full-length coding region for EmBP-1, a short stretch of DNA containing a stretch of T residues begins the sequence, suggesting the reversed fusion of a second cDNA (designated here as whGRP-1.1). When the 360 bp 3′ end fragment (nt 578–937 in Fig. 1A) was used as a hybridization probe to rescreen the library, 150 clones which hybridized strongly to the probe were identified from ca. 200000 phage on 5 plates, indicating that these clones are represented in the library at a much higher frequency (1:1333) than EmBP-1 cDNAs isolated from the same library which occurred at a frequency of 1:80000 [12, 22]. Preliminary northern blot data using this probe also agreed with this observation, resulting in very strong hybridization signals at a size distinct from that of EmBP-1 (not shown).

One of the clones (pXN14) appeared to be near full-length clone (nt 125–965 in Fig. 1A). A second screen was performed using a probe containing the 5′ end cDNA sequence from 125 (5′ end of pXN14) to 315 (a *Sac*I site in pXN14), and 150 additional positive clones were identified from another 200000 phage. From these two screens, ca. 40 phage isolates were purified and restriction mapped, and of these, 10 were selected for further analysis. Large-scale preparation of purified phage by CsCl equilibrium centrifugation and subcloning of DNA inserts from λgt11 phage DNA into pUC19 plasmid were performed by standard methods [27]. DNA sequencing of two near full-length clones (pXN28 1–901 and pXN14 125–965 in Fig. 1A), which were identical over their 778 bp overlap, and of 10 partial clones which were also identical in their overlaps resulted in the full-length sequence of whGRP-1 (Fig. 1A). The whGRP-1 sequence is 965 bp long with a poly(A) tail of 28 bp. Based on the location of the longest open reading frame, 5′- and 3′-untranslated regions of 100 and 330 bp, respectively, can be identified.

The longest open reading frame within this cDNA (101–604) encodes a protein (designated as whGRP-1) which shares strong sequence similarity to a family of RNA binding proteins, the RNPs [5]. The predicted whGRP-1 protein contains a RNP motif, characteristic of this family of RNA-binding proteins found in organisms ranging from bacteria to humans [5] (Fig. 1B). Within the RNP motif, the two conserved sub-domains RNP-1 (octamer) and RNP-2 (hexamer) were also found in the wheat cDNA. The molecular mass of whGRP-1 is predicted to be 16 kDa with a pI of 4.9. whGRP-1 also encodes a domain very rich in glycine residues (Fig. 1), with stretches of glycines interspersed with basic amino acids such as arginine, a characteristic shared with a large group of plant glycine-rich RNA-binding proteins (GRPs). Sequence alignment of 14 plant GRPs indicates that RNA-binding domains (residues 1–86) among these proteins isolated from monocots and dicots are highly conserved, their amino acid sequence similarities ranging from 69% to 94% (Fig. 1B). They all have glycine-rich domains

```
CTCCTCCCCTTCCGGTTCGGGTTCTGGTTCGGTTCGGGTCCGAGAGTAGCTAGCCTAGTC 60
TAGGGTTTCCGTAGAGAGGAAGGGAGAGGGAGATCCATCCATGGCGGAGACGGAGTACCG 120
                                          M  A  E  T  E  Y  R
CTGCTTCGTGGGCGGCCTCGCCTGGGCCACCGACGACAACAACCTCCAGCAGGCCTTCAG 180
 C  F  V  G  G  L  A  W  A  T  D  D  N  N  L  Q  Q  A  F  S
       RNP2
CCAGTACGGCGAGATCCTCGACGCCAAGATCATCAACGACCGCGAGACGGGGAGGTCCCG 240
  Q  Y  G  E  I  L  D  A  K  I  I  N  D  R  E  T  G  R  S  R
CGGGTTCGGCTTCGTCACGTTCGGCAGCGAGGAGTCGATGCGCCAGGCCATCGAGGAGAT 300
  G  F  G  F  V  T  F  G  S  E  E  S  M  R  Q  A  I  E  E  M
          RNP1
GAACGGCAAGGAGCTCGACGGGCGCAACATCACCGTCAACGAGGCCCAGTCCCGCCGCTC 360
  N  G  K  E  L  D  G  R  N  I  T  V  N  E  A  Q  S  R  R  S
GGGCGGAGGGGGCGGCGGCGGCGGCGGCTACGGCGGCCAGCGCGGCGGTGGCGGAGGCTA 420
  G  G  G  G  G  G  G  G  Y  G  G  Q  R  G  G  G  G  G  Y
CGGCGGCGGCGGCGGCTACGGAGGCGGCGGCGGCGGCTACGGCGGCCAGGGCGGTGGCGG 480
  G  G  G  G  G  Y  G  G  G  G  G  Y  G  G  Q  G  G  G  G
CTACGGCGGCCAGCGCGGTGGCGGCGGCGGCTACGGCGGCGGCGGCGGCTACGGCGGTGG 540
  Y  G  G  Q  R  G  G  G  G  G  Y  G  G  G  G  G  Y  G  G  G
TGGCGGCTACGGCGGCCAGCGCGGCGGTGGCGGCGGCGACTCCGGCGGCCAGTGGAGGAA 600
  G  G  Y  G  G  Q  R  G  G  G  G  G  D  S  G  G  Q  W  R  N
CTGAGCCTCGAATCTCCCAGTTATCGATCTATCTACCGTGTCGTCGTCGTCGTCTTCGTT 660
 •
AGCAAGTTATCTTCGGTGTCGTCGTGTTCCTGTGTCAGTCTGAGTCGTCGTGTGTGTGGT 720
GTTTTTGATCGCGAGCTCGAGTGTCGCTCTGTCCGGGGTTAATTAGAGCTGGTTACGCGA 780
GGAAGAAACAACATGTGTTACTATCTTTCAGTCGGTCGGTCGGTCGGTCAGTCGGTGATG 840
CAGCAATCCCCTTCTGTCATGGATTGGGATGTGATGTTGGAATATCCTATGTTTGAACAA 900
GTTTAAATGGATGAAGAAAATGCAAGTCATAGCTGCGAAAAAAAAAAAAAAAAAAAAAAA 960
AAAAA                                                          965
```

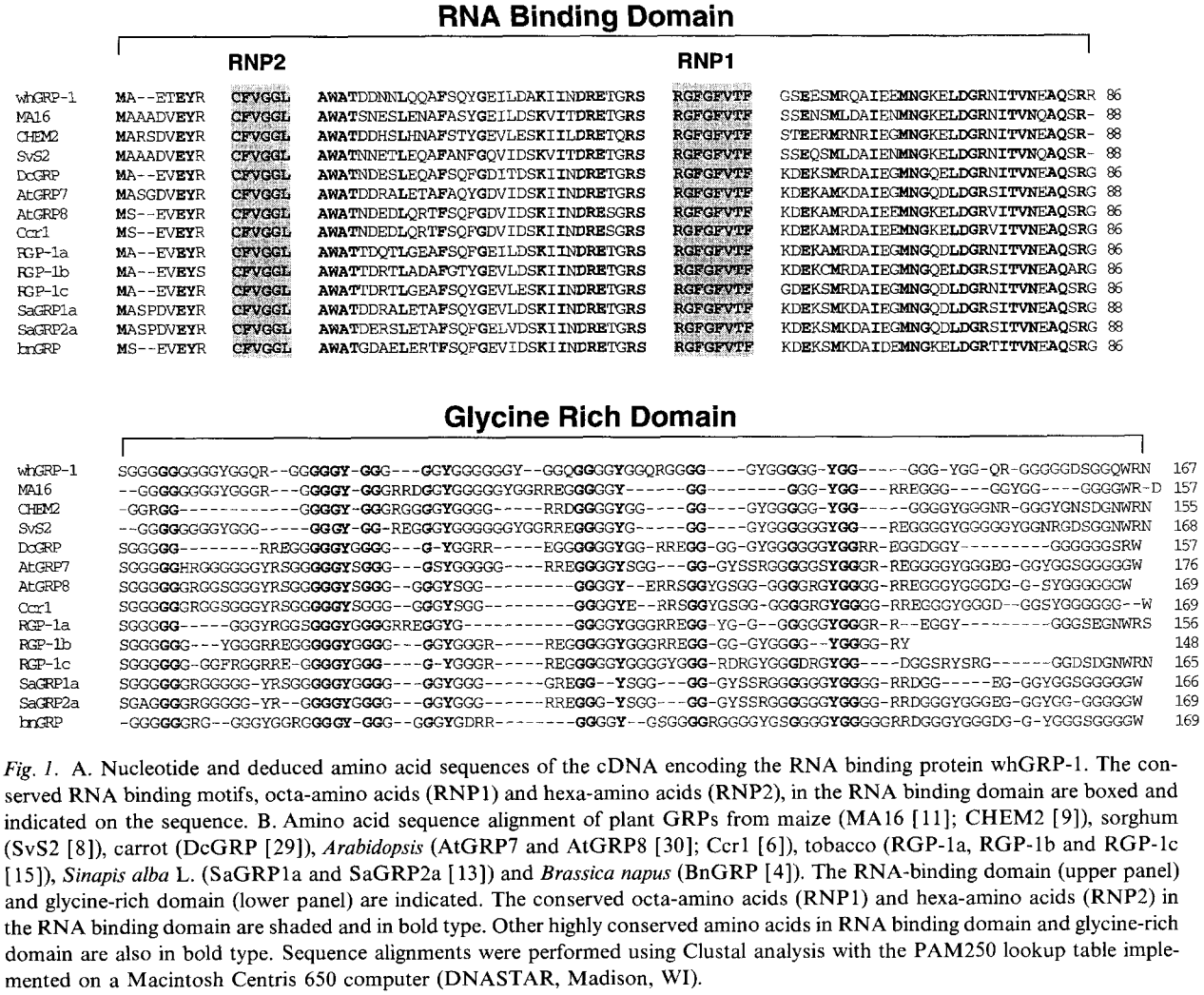


*Fig. 1.* A. Nucleotide and deduced amino acid sequences of the cDNA encoding the RNA binding protein whGRP-1. The conserved RNA binding motifs, octa-amino acids (RNP1) and hexa-amino acids (RNP2), in the RNA binding domain are boxed and indicated on the sequence. B. Amino acid sequence alignment of plant GRPs from maize (MA16 [11]; CHEM2 [9]), sorghum (SvS2 [8]), carrot (DcGRP [29]), *Arabidopsis* (AtGRP7 and AtGRP8 [30]; Ccr1 [6]), tobacco (RGP-1a, RGP-1b and RGP-1c [15]), *Sinapis alba* L. (SaGRP1a and SaGRP2a [13]) and *Brassica napus* (BnGRP [4]). The RNA-binding domain (upper panel) and glycine-rich domain (lower panel) are indicated. The conserved octa-amino acids (RNP1) and hexa-amino acids (RNP2) in the RNA binding domain are shaded and in bold type. Other highly conserved amino acids in RNA binding domain and glycine-rich domain are also in bold type. Sequence alignments were performed using Clustal analysis with the PAM250 lookup table implemented on a Macintosh Centris 650 computer (DNASTAR, Madison, WI).

and approximately uniform size (148 to 176 amino acids). Detailed examination indicates that whGRP-1 differs from all other plant GRPs in several positions; threonine (T) at position 4 and arginine (R) at position 86. whGRP-1 is also unique in containing several glutamine (Q) residues in the glycine-rich domain, a characteristic not found in any other plant GRPs identified to date. Whether these variations result in biochemical and biological differences remains unknown. Of the plant GRPs, whGRP-1 is most similar to maize CHEM2 and MA16 proteins and to sorghum SvS2, the only other monocot GRPs in the database.

*Northern blot analysis*

Expression of the whGRP-1 and Em genes in wheat leaves, roots, embryos and seedlings germinated with or without ABA as detected by northern blot analysis is shown in Fig. 2. Wheat total RNA was isolated by phenol extraction by a standard method [1] from plants which were grown to maturity in a greenhouse, and from developing wheat embryos excised and staged as previously described [26]. Root and leaf samples were excised from seedlings imbibed on MS salts on filter papers in a glass dish incubated in a growth room with lights for 5 days. For ABA treatment, wheat seeds were incubated in the dark for 3 days on filter papers soaked with MS salts with or without 0.1 mM ABA. As judged by hybridization to a 16S rRNA probe, the RNA amounts loaded in each lane were approximately the same.

The whGRP-1 mRNA steady state levels are relatively low in leaves as compared to the relatively high expression levels in roots and embryos.

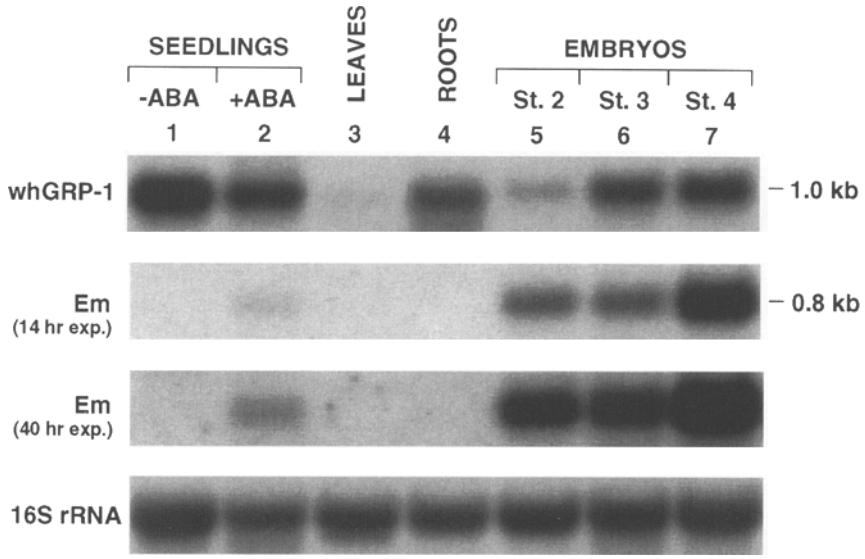


*Fig. 2.* Northern analysis of whGRP-1 gene transcripts in different tissues or at different development stages as indicated at the top of each lane. Approximately 10 μg of total RNA from each tissue were separated on a 1.2% agarose-formaldehyde gel and transferred to a nylon membrane [1]. Lanes 1 and 2, seedlings without and with ABA induction; lane 3, leaves; lane 4, roots; lanes 5–7, stage 2, 3, and 4 of embryos. As controls, Em gene transcripts and 16S rRNA on the same northern blot were also analyzed. A longer exposure (40 h exposure) is to show ABA induction of Em gene expression in seedlings. DNA probes were; a 320 bp *Sac*I-digested whGRP-1 cDNA fragment, a 680 bp *Pst*I-digested Em cDNA fragment [18], and a 400 bp *Eco*RI-digested 16S tomato ribosomal cDNA fragment which were gel-purified and labeled by the random primed method. Hybridization was in 5 × SSC (1 × SSC is 0.15 M NaCl, 0.015 M sodium citrate pH 7.0), 5 × Denhardt's solution, 0.5% (w/v) SDS, 100 μg/ml denatured salmon sperm DNA with a [$\alpha$-$^{32}$P]dCTP-labeled probe at 68 °C for 16 h. The filters were washed twice in 2 × SSC, 0.1% (w/v) SDS at room temperature and twice in 0.1 × SSC, 0.1% (w/v) SDS at 68 °C for 30 min each. Between probes, the filter was stripped by immersion in near-boiling 0.1% (w/v) SDS followed by cooling to room temperature and air-drying before rehybridization to a second probe as described above. Autoradiography was performed at −80 °C with intensifying screens and Kodak XAR 5 film.

Although we have not quantitatively determined the percentage of mRNA represented by the whGRP-1 transcript, judging from the very high signal strengths (exposure times of 7 h in Fig. 2), whGRP-1 hybridizing mRNA appears to be very abundant in these tissues. Our results for whGRP-1 expression in leaves and roots are consistent with the expression patterns of several other plant GRP genes [11, 15].

*Expression of whGRP in seedlings in not ABA-induced*

The maize MA16 gene (which is closely related to whGRP) was shown to be ABA inducible in embryos [11]. It has been proposed that MA16 may function in some aspect of ABA regulated gene expression. Because of the high similarity in sequence between MA16 and whGRP-1 (Fig. 1B), we tested if the whGRP-1 gene can be similarly induced by exogenous ABA in young seedlings. As a positive control for ABA induction, the well characterized Em gene [3, 25] was used as a second hybridization probe. The northern blot depicted in Fig. 2 reveals that the whGRP-1 gene is highly expressed in seedlings, however, exogenous ABA had no or a small repressive effect on whGRP-1 mRNA accumulation (compare lane 1 and lane 2 in Fig. 2). As expected, Em mRNA accumulation was induced by exogenous ABA applied to the seedlings [3], although the level of Em message was apparently much lower than that of the whGRP gene.

*Expression of whGRP-1 during wheat embryogenesis*

The whGRP cDNA was isolated from a cDNA library made from ABA treated wheat embryos, and thus embryo development presented another opportunity to test the potential ABA inducibility of the whGRP-1 gene. Northern blots of RNA extracted from dissected embryos at several stages of development [26] were hybridized to the whGRP-1 and Em probes (Fig. 2, lanes 5 to 7). Steady-state mRNA levels of both the whGRP-1 and Em transcripts are at relatively low levels early in development and accumulate to very high levels during later stages, but the timing of the increase in expression levels differs between the genes. A sharp increase in whGRP-1 transcripts can be observed between stage 2 and 3 embryos, prior to the known period of rapid ABA accumulation and induction of the maturation pathway [24]. As expected, although detectable in stage 2 embryos, a sharp increase in Em transcripts occurs later, during the transition from stage 3 to 4 of embryo development consistent with previous results [21, 31].

From these data it appears that the whGRP gene is regulated in a manner quite different from the ABA-inducible Em gene both in tissue specificity and during embryo development. The whGRP gene does not appear to be regulated by ABA in seedlings, and reaches maximal expression in developing embryos prior to the large increase in ABA levels during the transition to the maturation pathway.

## Acknowledgements

We thank Dr Mike Parker for assistance in DNA sequencing and analysis, Dr Ralph Quatrano and his lab for providing blotted RNA filters and Dr Rick Bostock who was involved in the original screen resulting in the fortuitous isolation of whGRP-1.1. This work was supported in part by grants to M.J.G. from the NSF (MCB 920 6095) and from The Pennsylvania State University, College of Agricultural Sciences (CRIS 3278).

## References

1. Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K: Current Protocols in Molecular Biology. Greene Publishing Associates/Wiley Interscience, New York (1987).
2. Bar-Zvi D, Shagan T, Schindler U, Cashmore AR: RNP-T, a ribonucleoprotein from Arabidopsis thaliana, contains two RNP-80 motifs and a novel acidic repeat arranged in an α-helix conformation. Plant Mol Biol 20: 833–838 (1992).

3. Berge SK, Bartholomew DM, Quatrano RS: Control of the expression of wheat embryo genes by abscisic acid. In: Goldberg RL (ed) Molecular Basis of Plant Development, pp. 193–201. Alan R. Liss, New York (1989).
4. Bergeron D, Beauseigle D, Bellemare G: Sequence and expression of a gene encoding a protein with RNA-binding and glycine-rich domains in *Brassica napus*. Biochim Biophys Acta 1216: 123–125 (1993).
5. Burd CG, Dreyfuss G: Conserved structures and diversity of functions of RNA-binding proteins. Science 265: 615–621 (1994).
6. Carpenter CD, Kreps JA, Simon AE: Genes encoding glycine-rich *Arabidopsis thaliana* proteins with RNA-binding motifs are influenced by cold treatment and an endogenous circadian rhythm. Plant Physiol 104: 1015–1025 (1994).
7. Cook WB, Walker JC: Identification of a maize nucleic acid-binding protein (NBP) belonging to a family of nuclear-encoded chloroplast proteins. Nucl Acids Res 20: 359–364 (1992).
8. Cretin C, Puigdomenech P: Glycine-rich RNA-binding proteins from *Sorghum vulgare*. Plant Mol Biol 15: 783–785 (1990).
9. Didierjean L, Frendo P, Burkard G: Stress responses in maize: sequence analysis of cDNAs encoding glycine-rich proteins. Plant Mol Biol 18: 847–849 (1992).
10. Dreyfuss G, Matunis MJ, Pinol-Roma S, Burd CG: hnRNP proteins and the biogenesis of mRNA. Annu Rev Biochem 62: 289–321 (1993).
11. Gomez J, Sanchez-Martinez D, Stiefel V, Rigau J, Puigdomenech P, Pages M: A gene induced by the plant hormone abscisic acid in response to water stress encodes a glycine-rich protein. Nature 334: 262–264 (1988).
12. Guiltinan MJ, Marcotte WR, Quatrano RS: A plant leucine protein that recognizes an abscisic acid response element. Science 250: 267–271 (1990).
13. Heintzen C, Melzer S, Fischer R, Kappeler S, Apel K, Staiger D: A light- and temperature-entrained circadian clock controls expression of transcripts encoding nuclear proteins with homology to RNA-binding proteins in meristematic tissue. Plant J 5: 799–813 (1994).
14. Higgins CF: Stability and degradation of mRNA. Curr Opin Cell Biol 3: 1013–1018 (1991).
15. Hirose T, Sugita M, Sugiura M: cDNA structure, expression and nucleic acid-binding properties of three RNA-binding proteins in tobacco: occurrence of tissue-specific alternative splicing. Nucl Acids Res 21: 3981–3987 (1993).
16. Hirose T, Suigita M, Sugiura M: Characterization of a cDNA encoding a novel type of RNA-binding protein in tobacco: its expression and nucleic acid-binding properties. Mol Gen Genet 244: 360–366 (1994).
17. Li Y, Sugiura M: Three distinct ribonucleoproteins from tobacco chloroplasts: each contains a unique amino terminal acidic domain and two ribonucleoprotein consensus motifs. EMBO J 9: 3059–3066 (1990).
18. Litts JC, Colwell GW, Chakerian RL, Quatrano RS: The nucleotide sequence of a cDNA clone encoding the wheat Em protein. Nucl Acids Res 15: 3607–3618 (1987).
19. Mattaj IW: RNA recognition: a family matter? Cell 73: 837–840 (1993).
20. Mieszczak M, Klahre U, Levy JH, Goodall GJ, Filipowicz W: Multiple plant RNA binding proteins identified by PCR: expression of cDNAs encoding RNA binding proteins targeted to chloroplasts in *Nicotiana plumbaginifolia*. Mol Gen Genet 234: 390–400 (1992).
21. Morris PC, Kumar A, Bowles DJ, Cuming AC: Osmotic stress and abscisic acid induce expression of the wheat Em genes. Eur J Biochem 190: 625–630 (1990).
22. Niu X, Guiltinan MJ: DNA binding specificity of the wheat bZIP protein EmBP-1. Nucl Acids Res 22: 4969–4978 (1994).
23. Ohta M, Sugita M, Sugiura M: Three types of nuclear genes encoding chloroplast RNA-binding proteins (cp29, cp31 and cp33) are present in *Arabidopsis thaliana*: presence of cp31 in chloroplasts and its homologue in nuclei/cytoplasms. Plant Mol Biol 27: 529–539 (1995).
24. Quatrano RS: The role of hormones during seed development. In: Davies PJ (ed) Plant Hormones and their Role in Plant Growth and Development, pp. 494–514. Martinus Nijhoff, Dordrecht, The Netherlands (1987).
25. Quatrano RS, Guiltinan MJ, Marcotte WR: Regulation of gene expression by abscisic acid. In: Verma DPS (ed) The Control of Gene Expression, pp. 69–90. CRC Press, Boca Raton, FL (1992).
26. Rogers SO, Quatrano RS: Morphological staging of wheat caryopsis development. Am J Bot 70: 308–311 (1983).
27. Sambrook J, Fritsch EF, Maniatis T: Molecular Cloning: A Laboratory Manual, Ed 2. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989).
28. Schuster G, Gruissem W: Chloroplast mRNA 3′ end processing requires a nuclear-encoded RNA-binding protein. EMBO J 10: 1493–1502 (1991).
29. Sturm A: A wound-inducible glycine-rich protein from *Daucus carota* with homology to single-stranded nucleic acid-binding proteins. Plant Physiol 99: 1689–1692 (1992).
30. van Nocker S, Vierstra RD: Two cDNAs from *Arabidopsis thaliana* encode putative RNA binding proteins containing glycine-rich domains. Plant Mol Biol 21: 695–699 (1993).
31. Williamson JD, Quatrano RS, Cuming AC: Em polypeptide and its messenger RNA levels are modulated by abscisic acid during embryogenesis in wheat. Eur J Biochem 152: 501–507 (1985).
32. Ye L, Li Y, Fukami-Kobayashi K, Go M, Konishi T, Watanabe A, Sugiura M: Diversity of a ribonucleoprotein family in tobacco chloroplasts: two new chloroplast ribonucleoproteins and a phylogenetic tree of ten chloroplast RNA-binding domains. Nucl Acids Res 19: 6485–6490 (1991).