

## Update section

### Sequence

# Nucleotide sequence of the rice (*Oryza sativa*) Em protein gene (*Emp1*)

J.C. Litts<sup>1</sup>, M.B. Erdman<sup>2</sup>, N. Huang, E.E. Karrer, A. Noueir<sup>3</sup>, R.S. Quatrano<sup>4</sup> and R.L. Rodriguez\*  
*Department of Genetics, University of California, Davis, CA 95616;* <sup>4</sup>*Department of Biology, University of North Carolina, Chapel Hill, NC 27599-3280, USA (\* author for correspondence); present addresses:* <sup>1</sup>*Cetus Corporation, 1400 53rd St., Emeryville, CA 94608, USA;* <sup>2</sup>*Cytogenetics Laboratory, UC Davis Medical Center, Sacramento, CA 95817, USA;* <sup>3</sup>*Department of Plant Pathology, University of California, Davis, CA 95616, USA*

Received 3 January 1992; accepted in revised form 22 January 1992

**Key words:** desiccation tolerance, osmotic stress, abscisic acid-regulated gene, seed maturation

The Em protein was first identified as a major product of the *in vitro* translation of poly(A) RNA obtained from dry wheat embryos [3, 4]. It was later shown that Em protein accumulation begins late in embryogenesis and that the level of accumulation can be manipulated in culture by abscisic acid (ABA) [20]. The wheat Em protein and its encoding mRNA are rapidly degraded during seed germination under normal conditions [2, 4, 18]. This decline is inhibited by ABA. Em transcripts also accumulate in wheat seedlings when treated with exogenous ABA or when subjected to water stress [1, 13]. This correlation between Em accumulation and water stress, along with the interesting physical properties of the Em protein, suggests that Em, along with other factors in the cell, may act as a cytoplasm protectant during desiccation [4, 12]. A cDNA clone (p1015) [8] and a genomic counterpart ( $\lambda$ CS41) [9] have been isolated and characterized from wheat. The 5' flanking region of the wheat gene represented by  $\lambda$ CS41 has been shown to contain elements that are responsive to ABA in a rice protoplast transient expression system [10] and in transgenic

tobacco [11]. A protein which binds one of these elements has been identified and a cDNA representing this protein has been characterized [5].

A genomic clone  $\lambda$ OSg4B containing a rice Em gene was isolated from an EMBL3 library [7] using a fragment of the wheat Em clone  $\lambda$ CS41 [9] as a heterologous probe. The nucleotide sequence of the rice Em gene (*Emp1*) is presented in Fig. 1. Southern blot analysis and trisomic mapping experiments have shown that the Em sequence in this clone is single copy in the rice genome and maps to chromosome 5 [16]. The nucleotide sequence includes 733 bp of 5' flanking region above the putative transcription start site, 114 bp of 5' untranslated leader, 395 bp of sequence including the coding region and a 107 bp intron, and 683 bp of 3' flanking region. A comparison of the *Emp1* nucleotide sequence with the sequences of several genes that are similar with respect to their products and/or regulation revealed several interesting similarities. The most interesting of these sequences was found in *Emp1* as a 15 bp perfect, inverted repeat located at 574 and inverted at 654. A similar sequence was found

```

1 GAGCCTCTAGGCTA GAAGAAATCCTAACG CCACCCTAATCTGA AGCAGGAAACAAATC AATTAAAGAAGCCAT
76 AAATGATAAATGAAT AATCCTTCATGCATA AAAATATTTAGTTAT CCATAAAGAGACGAC AGACGTATGCAACGT
151 TCATATTAATTGTCA CTCTACAGTGTCAAC GAGTTGTGGTCCAA GCAATTTTACCATA CATGAAAACCATCTG
226 AGAAAGGGCATGAGA GTAGTACTGCGTCAA TACACATGCAGTTGC ACACAGTAGCTGAGA GAGCGAGCCGGAAC
301 ATCTGGTCGCCGGTC CACGATCGGTCTACG CATGGGCTTCCCCCC GACCAGGTGGTGC GCGAGCCTACCACAT
376 GTCCACATGCATGCA CGGTGATCGACATCC CGATCGACGTGTGCA CAGTGCAGCCACCCC GCCGCTCGCTCGCTC
451 GCGCTCGCGAGTCCG GCTAGCACGGACATA CGGAGTAGTATATGC GTATGCATAGGCACG CGCGCGCACCGCGCA
526 TGCAAGCAACCAGAG CGCGCCGCGCACGGG CGGAGGCATAGCTTG CGCACGTACGTGTCCG CGCTCCGGCGCCCTC
601 GCCACGACGCGTGTG GCGGAGGCGCGCGCG CGCACCCGAGCCGTCC GCGTCCGCGCGCGAC ACGTACGTGCGCGC
676 GCCCTGATCCCGCCG CGGAGTGCGCCCTA TAAGAAGGACAGCGC GGCCGCGTCTCCAT CCCCATCAGCTCAAG
751 CCGCAGAAGACATAC ACACACAACAACAAG CCACCCTTCCGATTT GTTCATCGATCAGTT CGCAGCGTACGTGAC
826 GCTAGCTAACTAGTG TTTGGCA ATG GCG TCC GGG CAG CAG CAG CAG GGC AGG TCG GAG CTG GAC
      M A S G Q Q Q Q G R S E L D
890 CGC ATG GCC AGG GAG GGC CAG ACC GTC GTC CCC GGC GGC ACC GGC GGC AAG AGC CTC GAG
      R M A R E G Q T V V P G G T G G K S L E
950 GCC CAG GAG AAC CTC GCC GAG G | GTATGCAAAATAATA CTGAAACTTTTGATA GATCAAACCTGCCAT
      A Q E N L A E
1017 TTCATTGCGTATTTT GGTGAAGTACGTAA CTAATATGCGTACGT GTGCATCGATCGATC AG | GG CGC AGC
      G R S
1087 CGC GGG GGG CAG ACG AGG AAG GAG CAG ATG GGG GAG GAA GGG TAC CGC GAG ATG GGG CGC
      R G G Q T R K E Q M G E E G Y R E M G R
1147 AAG GGC GGC CTC AGC ACC GGC GAC GAG TCC GGC GGC GAG CGC GCC GCC CGC GAG GGC ATC
      K G G L S T G D E S G G E R A A R E G I
1207 GAC ATC GAC GAG TCC AAG TAC AAG ACC AAG TCC TAG ACTACACACACTTTT GCATCCGTGAATG
      D I D E S K Y K T K S Stop
1271 CCAGTGTGTCGTAGT CGTCTCAGTTGTAGT CATCGAGTCAGTCTT AGCTAGCTAGCTCTC TTATCAATAATAATG
1346 TAGGTCTTGACGGAT GCACGCATTTAGGCG CCCGTATGATTTGCT ATGTTATGTTTCATA TGGTATGCGTGTCTT
1421 AGCTTTAGCTAGCTT TGGTAGTTTGGTTTA GGTGCTGGTCACTAG CTAGGGTATGTGTTT TAGTTTATTGCAGGA
1496 TCGGTACGGCAGGGG CTGAATGCTAGCTAG GTGATGCTTTGGCTT TTCATGGCCAGCAGC TGCTAGCTTCTTGT
1571 TCCAACCTGATGCCTC TGCTGGGTGTGCTCT TTGTAACCGTGCTTG CTTTCAGTTAATCTA GCTAGCTGTTGTCA
1646 TCCGAAAGTGCCAGT GTCTATGTCTGATCT ATCGCGGCTAATTGT TGCTGTTGCTAGTTG ATCTTTGATTCCTCA
1721 GAGATGCCATGTGCA GGATGGAATTATTAG CGCAAAACACAGGAG AAGGACAACAATTT TAAGTTGCCTAATTT
1796 TTTTCTTAAACACC TGCTGGGCCGTGCAC ACAAATAAATAATCA AGTACCTCGTCTAAA CTGAATGCTCAAGTT
1871 TGAAAGTGAAAACAG GACGTGCCGTGCCAA GAACTTCGAGATATC TCCGAGTTCT

```

Fig. 1. Nucleotide sequence of *Empl* as determined from  $\lambda$ OSg4B. The coding region is blocked into codons and translated. The introns and exons are demarcated by vertical bars (|). The various putative control elements are underlined, with the exception of the ABREs which are double-underlined.

at 412. This repeat, ACGTACGTGTCGCGC, contains a sequence (underlined) which is very similar to those identified as ABA-response elements (ABREs) in a wheat *Em* gene [9,11] and another ABA-responsive rice gene [14,17]. ABREs and their role in gene regulation have been reviewed recently [6, 15].

An analysis of *Empl* expression in mature rice embryos revealed that it is regulated in a manner similar to that described in detail for the wheat

gene [19, 20]. When isolated mature rice embryos were incubated on moistened filter paper there was an 80% reduction in the amount of accumulated mRNA in the first six hours of germination. When similar embryos were incubated for 24 h in the presence of ABA, the level of *Empl* mRNA was maintained. Like the water control, the levels of mRNA in GA-treated embryos dropped by 80%. Rice *Empl* mRNA was also found to accumulate in immature seeds.

In summary, based on sequence analysis and preliminary expression studies, the rice *Empl* gene belongs to a highly conserved family of genes present in both monocot and dicot plants. *Empl* is the only gene in the rice genome closely related to the Em protein family. *Empl* contains highly conserved ABREs that are probably responsible for its induction by ABA, as well as other sequences that are conserved in similar genes and which may control its expression in other ways. *Empl* is regulated very similarly, if not identically, to the wheat Em genes.

### Acknowledgements

This work was supported in part by a gift from DuPont de Nemours & Co., by a grant to R.L.R. (USDA-CRGO-88-37262-4019) and by a grant to R.S.Q. (USDA-CRGO-89-37262-4456).

### References

1. Berge SK, Batholomew DM, Quatrano RS: Control of the expression of wheat embryo genes by abscisic acid. In R. Goldberg (ed) *The Molecular Basis of Plant Development*, pp. 193–201. Alan R. Liss, New York (1989).
2. Cuming AC: Developmental regulation of gene expression in wheat embryos. Molecular cloning of a DNA sequence encoding the early-methionine-labelled (Em) polypeptide. *Eur J Biochem* 145: 351–357 (1984).
3. Cuming AC, Lane BG: Protein synthesis in imbibing wheat embryos. *Eur J Biochem* 99: 217–224 (1979).
4. Grzelczak ZF, Sattolo MH, Hanely-Bowdoin LK, Kennedy TD, Lane BG: Synthesis and turnover of proteins and mRNA in germinating wheat embryos. *Can J Biochem* 60: 389–397 (1982).
5. Guiltinan MJ, Marcotte J, William R, Quatrano RS: A plant leucine zipper protein that recognizes an abscisic acid response element. *Science* 250: 267–271 (1990).
6. Hetherington AH, Quatrano RS: Mechanisms of action of abscisic acid at the cellular level. *New Phytol* 119: 9–32 (1991).
7. Huang N, Sutliff TD, Litts JC, Rodriguez RL: Classification and characterization of the rice  $\alpha$ -amylase multi-gene family. *Plant Mol Biol* 14: 655–668 (1990).
8. Litts JC, Colwell GC, Chakerian RL, Quatrano RS: The nucleotide sequence of a cDNA encoding the wheat Em protein. *Nucl Acids Res* 15: 3607–3618 (1987).
9. Litts JC, Colwell GW, Chakerian RL, Quatrano RS: Sequence analysis of a functional member of the Em gene family from wheat. *DNA Sequence* 1: 263–274 (1991).
10. Marcotte JWR, Bayley CC, Quatrano RS: Regulation of a wheat promoter by abscisic acid in rice protoplasts. *Nature* 335: 454–457 (1988).
11. Marcotte JWR, Russell SH, Quatrano RS: Abscisic acid-responsive sequences from the Em gene of wheat. *Plant Cell* 1: 969–976 (1989).
12. McCubbin WD, Kay CM: Hydrodynamic and optical properties of the wheat germ Em protein. *Can J Biochem* 68: 803–811 (1985).
13. 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 (1990).
14. Mundy J, Yamaguchi-Shinozaki K, Chua N-H: Nuclear proteins bind conserved elements in the abscisic acid-responsive promoter of a rice *rab* gene. *Proc Natl Acad Sci USA* 87: 1406–1410 (1990).
15. Quatrano RS, Guiltinan MJ, Marcotte JWR: Regulation of gene expression by abscisic acid. In: Verma D-PS (ed) *Control of Plant Gene Expression*, in press. Telford Press, Caldwell, NJ (1992).
16. Ranjhan S, Litts JC, Foolad MR, Rodriguez RL: Chromosomal localization and genomic organization of  $\alpha$ -amylase genes in rice (*Oryza sativa* L.). *Theor Appl Genet* 82: 481–488 (1991).
17. Skriver K, Mundy J: Gene expression in response to abscisic acid and osmotic stress. *Plant Cell* 2: 503–512 (1990).
18. Thompson EW, Lane BG: Relation of protein synthesis in imbibing wheat embryos to the cell-free translational capacities of bulk mRNA from dry and imbibing embryos. *J Biol Chem* 255: 5965–5970 (1980).
19. Williamson JD, Quatrano RS: ABA-regulation of two classes of embryo-specific sequences in mature wheat embryos. *Plant Physiol* 86: 208–215 (1988).
20. Williamson JD, Quatrano RS, Cuming AC: Em polypeptide and its messenger RNA levels are modulated by ABA during embryogenesis in wheat. *Eur J Biochem* 152: 501–507 (1985).