Plant Physiol. (1994) 106: 787–788

Plant Gene Register

A Third Highly Conserved Group 1 Lea Gene from *Arabidopsis thaliana* L.[1]

Éberson S. Calvo[2]*, Steven R. Rodermel, and Randy C. Shoemaker

Department of Zoology and Genetics (E.S.C., R.C.S.), Department of Botany (S.R.R.), United States Department of Agriculture-Agricultural Research Service, Field Crop Research (R.C.S.), and Department of Agronomy (R.C.S.), Iowa State University, Ames, Iowa 50011

Lea genes are thought to play a major role in protecting plant cells against desiccation (Dure et al., 1989). Although the function of the proteins encoded by these genes remains unclear, both developmental and environmental signals seem to be involved in triggering Lea gene expression in angiosperm embryos (Galau et al., 1991; Bostock and Quatrano, 1992). The hormone ABA may be involved in regulating Lea gene expression (Bostock and Quatrano, 1992), although its role as a developmental factor that regulates group 1 Lea gene expression during embryo development is still unclear (Galau et al., 1991; E.S. Calvo, E.S. Wurtele, and R.C. Shoemaker, unpublished data).

We have cloned a group 1 Lea gene from *Arabidopsis thaliana* ecotype Landsberg erecta by probing an *Arabidopsis* genomic library with a SLE cDNA sequence (E.S. Calvo, E.S. Wurtele, and R.C. Shoemaker, unpublished data). Nine genomic clones showing the strongest hybridization signals were restriction mapped and found to originate from the same locus, here referred to as *Ale* (*Arabidopsis* late embryogenic) (Table I). The *Ale* gene contains an open reading frame coding for a 112 amino acid polypeptide that shows extensive homology to the SLE, as well as to other group 1 LEA proteins; we have designated this protein as ALE.

A data base search in GenBank revealed that the cloned ALE sequence also shows high homology to the *D19h* gene from *A. thaliana* Landsberg, which encodes a 92 amino acid embryo abundant protein (Finkelstein, 1993). ALE shows 77% identity and 87% similarity with D19h protein. However, the ALE polypeptide is distinguished by the presence of a 20 amino acid imperfect repeat. Both genes contain a single intron at the same position but of variable length and sequence. Thus, it seems that two homologous loci for group 1 LEA proteins, differing mainly by the presence of a 20 amino acid repeated motif, are present in the genome of *A. thaliana* ecotype Landsberg erecta. A similar situation is found in the cotton genome (Espelund et al., 1992). Recently, two homologous genes, *GEA1* and *GEA6*, were cloned from *A. thaliana* ecotype Columbia (Gaubier et al., 1993). Although *GEA6* and *D19h* were found to encode identical polypeptides, *GEA1* and *Ale*-encoded polypeptides differ in the number of the 20 amino acid repeats, with the GEA1 polypeptide containing four repeats instead of two. Otherwise, these two genes appear to be identical at the nucleotide sequence level. Since Southern blot experiments at low-stringency conditions

Table I. *Characteristics of the Ale gene from Arabidopsis*

Organism:
Arabidopsis thaliana L. ecotype Landsberg erecta.
Function:
Unknown; putative role in protecting seeds against desiccation.
Cloning Techniques:
Screening of a λFIX *Arabidopsis* genomic library (Voytas et al., 1990) using a soybean LEA cDNA (SLE) as a heterologous probe. A 1.7-kb *SalI-PstI* fragment from one of the nine λ clones isolated in the screening was subcloned into pBluescript KS$^+$ (Stratagene, La Jolla, CA) and used to generate a unidirectional deletion series with exonuclease III (Promega, Madison, WI). Sequencing was performed on double-stranded templates using the dideoxy chain-termination method.
Expression and Regulation:
Undetermined; putative seed-specific expression during late embryo development.
Features of Gene Sequence:
A single intron, 187 bp in size. The 5′ untranscribed region contains a CACGT sequence found in the promoter of several ABA-responsive genes.
Gene Copy Number:
Probably one copy per haploid genome. A second locus (*D19h*) encoding a homologous group 1 Lea gene exists in the genome of *Arabidopsis*.
Amino Acid Sequence Identification:
An open reading frame of 112 amino acids was identified that shares extensive identity (80%) and similarity (88%) with the soybean SLE protein (E.S. Calvo, E.S. Wurtele, and R.C. Shoemaker, unpublished data) as well as with other members of group 1 LEA polypeptides (Dure et al., 1989). The ALE polypeptide has a 20 amino acid imperfect repeated motif also recently found in some members of group 1 LEA proteins.

[1] Joint contribution of North Central Region, U.S. Department of Agriculture-Agricultural Research Service and Journal Paper No. J-15743 of the Iowa Agriculture and Home Economics Experiment Station, Ames, IA 50011; Project No. 2974.

[2] E.S.C. was supported by a fellowship provided by the Brazilian National Council for Scientific Development.

* Corresponding author; fax 1–515–294–2299.

Abbreviations: Lea, late embryogenic abundant; SLE, soybean LEA.

detected only two hybridizing bands (data not shown; Finkelstein, 1993; Gaubier et al., 1993), we hypothesize that *GEA1* and *Ale* are allelic. The presence of this polymorphism between the two ecotypes should quickly allow the mapping of this gene with a PCR-based co-dominant marker into currently available *Arabidopsis* genetic maps. Furthermore, it may also suggest that the different group 1 LEA proteins containing different numbers of the 20 amino acid hydrophilic repeats (Espelund et al., 1992) are functionally equivalent in terms of the embryo's abilities to survive desiccation and germinate, since no major differences are apparent between the two *Arabidopsis* ecotypes.

The promoter region of the *Arabidopsis Ale* gene contains one conserved motif, a CACGT sequence at position 415, that may be involved in an ABA response pathway (Marcotte et al., 1989). However, no sequences similar to the "Sph element" (CATGCATG) appear to be present in the *Ale* promoter or in the promoter of its homologs. This sequence is involved in the Vp-1-mediated transcription activation of a seed-specific promoter in maize, and is also present in several other Lea genes (Hattori et al., 1992). Since the putative *Arabidopsis* Vp-1 homolog was recently isolated (Giraudat et al., 1992), the cloning of the *Arabidopsis Ale* gene will enable the design of experiments aimed toward elucidating the similarities and differences between the regulation of Lea genes in species such as maize versus *Arabidopsis*. Furthermore, it will be possible to use a genetic approach to isolate mutants in the transduction pathway that regulates both the developmental and environmental expression of Lea genes in *Arabidopsis*.

Received March 29, 1994; accepted April 10, 1994.

The GenBank/EMBL accession number for the sequence reported in this article is L16946.

LITERATURE CITED

Bostock RM, Quatrano RS (1992) Regulation of Em gene expression in rice. Plant Physiol **98:** 1356–1363

Dure L III, Crouch M, Harada J, Ho TD, Mundy J, Quatrano R, Thomas T, Sung ZR (1989) Common amino acid sequence domains among LEA proteins of higher plants. Plant Mol Biol **12:** 475–486

Espelund M, Stein S-L, Hughes W, Galau GA, Larsen F, Jakobsen KS (1992) Late embryogenesis-abundant genes encoding proteins with different numbers of hydrophilic repeats are regulated differentially by abscisic acid and osmotic stress. Plant J **2:** 241–252

Finkelstein RR (1993) Abscisic acid-insensitive mutations provide evidence for stage-specific signal pathways regulating expression of an *Arabidopsis* late embryogenesis-abundant (lea) gene. Mol Gen Genet **238:** 401–408

Galau GA, Jakobsen KS, Hughes W (1991) The controls of late dicot embryogenesis and early germination. Physiol Plant **81:** 280–288

Gaubier P, Raynal M, Hull G, Huestis GM, Grellet F, Arenas C, Pages M, Delseny M (1993) Two different Em-like genes are expressed in *Arabidopsis thaliana* seeds during maturation. Mol Gen Genet **238:** 409–418

Giraudat J, Hauge BM, Valon C, Smalle J, Parcy F, Goodman HM (1992) Isolation of the *Arabidopsis* ABI3 gene by positional cloning. Plant Cell **4:** 1251–1261

Hattori T, Vasil V, Rosenkrans L, Hannah LC, McCarty DR, Vasil IK (1992) The viviparous-1 gene and abscisic acid activate the C1 regulatory gene for anthocyanin biosynthesis during seed maturation in maize. Genes Dev **6:** 609–618

Marcotte WR Jr, Russell SH, Quatrano RS (1989) Abscisic acid-responsive sequences from the Em gene of wheat. Plant Cell **1:** 969–976

Voytas DF, Konieczny A, Cummings MP, Ausubel FM (1990) The structure, distribution and evolution of the Ta1 retrotransposable element family of *Arabidopsis thaliana*. Genetics **126:** 713–721