Plant Physiol. (1993) 101: 695–696

Plant Gene Register

Cotton *Lea5* and *Lea14* Encode Atypical Late Embryogenesis-Abundant Proteins[1]

Glenn A. Galau*, Helen Y.-C. Wang, and D. Wayne Hughes
Department of Botany, University of Georgia, Athens, Georgia 30602

LEA proteins comprise a large group of probable desiccation protectants that in cotton (*Gossypium hirsutum*) are developmentally induced during the postabscission stage of embryogenesis (Hughes and Galau, 1989) and are environmentally induced in embryos by desiccation or culture with ABA or high osmoticum (Hughes and Galau, 1991). In other plants, some *Lea* mRNAs are induced by similar stresses in vegetative tissues (Skriver and Mundy, 1991). Of the many *Lea* and water stress-related mRNAs that have been sequenced (Skriver and Mundy, 1991), all but cotton *Lea12*, cDNA D34 (Baker et al., 1988), and desiccation-induced *Craterostigma* cDNA pcC27-45 (Piatkowski et al., 1990) encode proteins that are very hydrophilic (Baker et al., 1988; Skriver and Mundy, 1991). Extensive hydrophilicity is expected of structural and nonstructural proteins that bind or replace water during water stress (Baker et al., 1988; Hughes and Galau, 1989), but other functions might be expected for those that are not remarkably hydrophilic. Of the 18 cloned cotton *Lea* and *LeaA* mRNAs (Galau et al., 1986), only *Lea5* (cDNA D73) and *Lea14* (cDNA D95) are highly induced in mature leaves of water-stressed plants or in water-stressed, detached leaves (D.W. Hughes, G.A. Galau, unpublished observations). We report here that both genes encode proteins with significant hydropathic character and that *Lea14* is a homolog of the gene encoding the *Craterostigma* desiccation-induced cDNA pcC27-45 (Piatkowski et al., 1990).

Both alloalleles of *Lea5* were recovered, and 3809 nucleotides of *Lea5-A* clone GD73-14R and 1269 nucleotides of *Lea5-D* clone GD73-2R were sequenced along with several *Lea5* cDNA clones (Table I). *Lea5-A* and *Lea5-D* are very similar from at least 240 nucleotides upstream of the transcription start to about 230 nucleotides 3′ of the major polyadenylation site, after which they diverge in sequence. The cDNA clones D73 and D96 are transcribed from *Lea5-A*, and cDNA clone D122 is transcribed from *Lea5-D*. Clone D73 is the result of an error in excision of the single intron; instead of splicing to *Lea5-A* nucleotide 2515, as occurs with D96 and at an equivalent position with *Lea5-D* cDNA D122, the D73 mRNA is spliced to *Lea5-A* nucleotide 2529, resulting in an out-of-frame translation of six amino acids before a premature termination. *Lea5-A* and *Lea5-D* encode 11.4-kD proteins with substitutions in only 6 of their 105 amino acids. Their hydropathic profiles are unremarkable; the amino-terminal half is somewhat hydropathic, possibly with a membrane-spanning region, and the carboxy-terminal half is somewhat hydrophilic.

The A genome alloallele of *Lea14* was recovered, and 4535 nucleotides of *Lea14-A* clone GD95-18 was sequenced along with cDNA clone D95. The sequence of a partial clone of the other alloallele, *Lea14-D* clone GD95-23B, contains several nucleotide substitutions in the carboxy-terminal portion of the encoded protein. The cDNA clone D95 is from the *Lea14-A* alloallele. *Lea14-A* encodes a 16.4-kD protein that is exactly colinear, with 66% identity, with that encoded by the *Craterostigma* cDNA pcC27-45, which is induced in leaves and roots during desiccation and in ABA-treated and NaCl-treated callus (Piatkowski et al., 1990). The proteins encoded by cotton *Lea14* and *Craterostigma* pcC27-45 thus define an additional family of water stress-related proteins (Dure et al., 1989), the Group 4 LEA proteins. These proteins are slightly hydropathic throughout.

An ACGT-containing element has been shown to be involved in the ABA induction of a wheat *Lea* gene (Guiltinan et al., 1990). *Lea14-A* contains sequences at nucleotides −58 and −14 from the transcription start that are similar to this element and similar sequences that are in many cotton *Lea* genes (Galau et al., 1992). However, *Lea5-A* and *Lea5-D* contains only one sequence, at nucleotide −127 from the transcription start, that is only partially similar to these.

Received August 28, 1992; accepted September 14, 1992.

The GenBank accession numbers for the sequences reported in this article are: *Lea5-A* GD73-14RXB144, M37697; *Lea5-D* GD73-2R, L01102; *Lea4-A* cDNA D96, M99324; *Lea5-D* cDNA D122, M88323; *Lea14-A* GD95-18XR, M88321; *Lea-14-A* cDNA D95, M88322.

[1] Supported by grants from the National Institutes of Health and the United States Department of Agriculture Competitive Grants Program.

* Corresponding author; fax 1–706–542–1805.

Abbreviation: LEA, late embryogenesis-abundant.

Table I. *Characteristics of Lea5 and Lea14 genes and cDNAs from Gossypium hirsutum*

Organism:
Gossypium hirsutum L. cv Coker 201 (Upland cotton), Malvaceae.
Function:
Encode probable desiccation protectants (Hughes and Galau, 1989, 1991).
Expression:
During the postabscission stage of embryo development and during water-related stress during embryo culture; coordinately expressed at these times with many other genes, including all *Lea*, *LeaA*, and *MatP* genes (Hughes and Galau, 1989, 1991).
Source:
Nuclear DNA from embryo cotyledons 20 to 23 d postanthesis (preendoreduplication). A partial *Sau*3AI digest was cloned in LambdaGEM-12 (Promega), and phage were identified by hybridization with cDNA clones D73 and D95. All cDNA clones have been described (Balau et al., 1986).
Lea5:
Lea5-A Clone GD73-14RXB144:
Two phage isolates contain the *Lea5-A* alloallele in the A genome, based on the size of the *Eco*RI and *Bam*HI fragments that hybridize to cDNA D73 (Galau et al., 1988). Cloning of the 6.5-kb *Eco*RI fragment of isolate 14 into Bluescript (Stratagene) was followed by subcloning of restriction fragments and deletions and complete dideoxy sequencing of both strands with *Taq* polymerase using double-stranded templates.
Lea5-D Clone GD73-2R:
Cloning of the 6.1-kb *Eco*RI fragment of phage isolate 2 of the D genome alloallele and sequencing on both strands, as described above.
Lea5-A cDNAs D73 and D96 and *Lea5-D* cDNA D122:
Subcloning of the cDNA inserts into Bluescript and complete sequencing of both strands as described above, using internal primers and cloned restriction fragments and deletions. D122 contains a portion of its 3′ end in reverse orientation on its 5′ end (not reported).
Lea14:
Lea14-A Clone GD95-18XR:
Three phage contain the *Lea14-A* alloallele in the A genome, based on the size of the *Eco*RI fragment that hybridizes with cDNA D95 (Galau et al., 1988). Cloning of the entire insert of isolate 18 and subcloning of an internal *Eco*RV/*Xba*I fragment and an upstream overlapping *Eco*RI fragment were followed by sequencing of both strands as described above.
Lea14-A cDNA D95:
Subcloning of the *Pst*I + *Pvu*I insert (Balau et al., 1986) into Bluescript and sequencing of both strands as described above.
Transcription Starts:
Primer extension assays were used to map the transcription starts of mRNAs present in post-abscission-stage embryos. Primers were *Lea5-A* nucleotide 2358–2338 and *Lea14-A* nucleotide 2176–2158.
Structural Features of the Proteins:
Both *Lea5* and *Lea14* contain hydrophathic regions. *Lea14* is a homolog of the *Craterostigma* gene encoding the cDNA pcC27-45 (Piatkowski et al., 1990).

LITERATURE CITED

Baker J, Steele C, Dure L III (1988) Sequence and characterization of 6 *Lea* proteins and their genes from cotton. Plant Mol Biol **11:** 277–291

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

Galau GA, Bass HW, Hughes DW (1988) Restriction fragment length polymorphisms in diploid and allotetraploid *Gossypium*: assigning the late embryogenesis-abundant (*Lea*) alloalleles in *G. hirsutum*. Mol Gen Genet **211:** 305–314

Galau GA, Hughes DW, Dure L III (1986) Abscisic acid induction of cloned cotton late embryogenesis-abundant (*Lea*) mRNAs. Plant Mol Biol **7:** 155–170

Galau GA, Wang HY-C, Hughes DW (1992) Cotton *Lea4* (D19) and *LeaA2* (D132) Group 1 *Lea* genes encoding water stress-related proteins containing a 20-amino acid motif. Plant Physiol **99:** 783–788

Guiltinan MJ, Marcotte WR Jr, Quatrano RS (1990) A plant leucine zipper protein that recognizes an abscisic acid response element. Science **250:** 267–271, 612

Hughes DW, Galau GA (1989) Temporally modular gene expression during cotyledon development. Genes Dev **3:** 358–369

Hughes DW, Galau GA (1991) Developmental and environmental induction of *Lea* and *LeaA* mRNAs and the postabscission program during embryo culture. Plant Cell **3:** 605–618

Piatkowski D, Schneider K, Salamini F, Bartels D (1990) Characterization of five abscisic acid-responsive cDNA clones isolated from the desiccation-tolerant plant *Craterostigma plantagineum* and their relationship to other water-stress genes. Plant Physiol **94:** 1682–1688

Skriver K, Mundy J (1991) Gene expression in response to abscisic acid and osmotic stress. Plant Cell **2:** 503–512