Journal of Genetics and Molecular Biology
Vol. 19, No. 3, 168-181, September 1, 2008

Expression of *Glycine max* Physiologically Mature Genes in Soybean (*Glycine max* L.) Tissues

Yuan-Ching Tsai[1], Ming-Tsung Wu[1], Jaw-Shu Hsieh[2], Yue-Ie C. Hsing[1] and Ming-Der Shih[1*]

1. Institute of Plant and Microbial Biology, Academia Sinica, Taipei, 11529, Taiwan
2. Department of Agronomy, National Taiwan University, Taipei, 10617, Taiwan

Late embryogenesis abundant (LEA) genes, in abundance in seeds during the late stages of development, are associated with desiccation tolerance. In the present work, we characterized the expression patterns of 15 soybean Glycine max physiologically mature (GmPM) genes, including 11 typical and 4 atypical LEA genes, by RT-PCR and in silico analysis. The transcripts of all tested genes were detected in mid-developmental or pod-dried soybean seeds. However, only the expression of GmPM6, a LEA II gene, was significantly enhanced under high salt and dehydration conditions. In silico study confirmed this observation. Under normal conditions, GmPM4, 6, 17, 25, 30, and 39 were expressed in mature flowers. In addition, GmPM4, 5, 6, 17, 25, and 28 expression was detected in several vegetative tissues, such as nodules, roots, stems, leaves, or germinating cotyledons. Therefore, soybean GmPM genes might be involved in a complex regulation pathway for osmotic stress.

Key words: expression patterns, in silico analysis, late embryogenesis abundant, RT-PCR, soybean

Introduction

Late embryogenesis abundant (*LEA*) genes are highly expressed in seeds during late development. In addition, some *LEA* genes can be induced in vegetative tissues by osmotic stress, including salt treatment and desiccation, as well as with exogenous abscisic acid (ABA) treatment [reviewed in 1-4]. The genes have been studied extensively and were isolated from plants, fungi, microbes, and even animals [reviewed in 3-5]. LEA proteins are classified into at least 5 groups on the basis of their conserved sequences [3,4,6] or at least 6 groups according to the "Protein or Oligonucleotide Probability Profile," which shows over- or under-representation of particular amino

* Corresponding authors: Ming-Der Shih; Institute of Plant and Microbial Biology, Academia Sinica, 128 Academia Rd Sec 2, Taipei, Taiwan. Tel: 886-2-2787-1049 FAX: 886-2-2782-7954 <patlabor@gate.sinica.edu.tw>

JGMB<http://140.109.54.28/jgmb>

acids in protein sequences [5,7]. The primary structures and hydropathy studies suggested that the LEA I to IV proteins are hydrophilic, whereas the LEA V proteins are hydrophobic. Structural analysis suggested that hydrophilic LEA proteins are members of "natively unfolded proteins" and change their conformations when internal water molecules are removed [e.g. 8-11]. The 3-D structures of a D-95-type LEA V protein, Arabidopsis AtLEA7-1, revealed an αβ-fold consisting of one α-helix and seven β-strands that form two antiparallel β-sheets [12].

The expression profiles of *LEA* genes in diverse species were also reported. In general, most *LEA* genes, such as rice *Rab16A-D* [13], barley *HVA1* [14], soybean *GmPM8* [15] and tomato *Le25* [16], are expressed in both reproductive and vegetative organs. However, some, such as wheat *Em* [17], cotton *D19* [18], carrot *Dc8* [19], and *B. napus Lea76* [20], are seed or embryo specific, and others, such as Arabidopsis *Cor47* [21] and *Lit29* [22], spinach *Cap85* [23], and several *Craterostigma plantagineum LEA II*, *III*, *IV*, and *V* genes [24-26], are vegetative-tissue specific. Detailed analysis revealed that the expression of *LEA* genes is regulated significantly by temporal and spatial effects. The mRNA expression of *HVA1*, a barley *LEA III* gene, was induced in roots, as well as in coleoptiles and leaves, in ABA- or drought-treated 3-day-old seedlings but was barely detected in leaves and coleoptiles in 7-day-old treated seedlings [14]. The expression patterns of two Arabidopsis *Em* genes, *AtEm1* and *AtEm6*, showed *AtEm1* transcripts first detected about 13-14 days after pollination and mainly accumulated in vasculature and pericycle, whereas those of *AtEm6* were detected 3 days later and expressed throughout the embryo [27,28]. Therefore, considering their redundancy and high level of expression, extreme hydrophilic property, protein structure, as well as timing for accumulation, LEA proteins were suggested to be one of the key factors in the acquisition of desiccation tolerance. For example, the protein accumulation of barley HVA1 conferred water-deficit and salt-stress tolerance in transgenic cereals [29-32].

Orthodox seeds undergo dehydration and maturation during the late developmental stage. Several studies from various species suggested that the moisture loss during seed maturation is critical for the embryo phase transition [33,34]. Interestingly, artificially drying immature soybean seeds also revealed similar embryo phase transition. This process was termed as precociously maturation or physiology maturation [35]. We isolated 41 cDNA clones from precociously matured soybean seeds and characterized several of them [10,15,36-39]. These clones were designated *GmPM*, for *Glycine max* physiologically mature. Over half of them belong to *LEA* genes, including 1 *LEA I*, 3 *LEA II*, 6 *LEA III*, 6 *LEA IV*, and 5 *LEA V* and 7 atypical *LEA* genes. Although the abundance of *LEA* genes during treatment to induce precocious maturation suggests the relation between *LEA* genes and water loss and the expression of numerous *LEA* genes is induced under exogenous ABA or environmental stress treatment, little information exists on the expression profiles of soybean *LEA* genes under hormone or water-deficit treatment. In the present study, we characterized the expression patterns of 15 soybean *GmPM* genes, including 11 typical and 4 atypical *LEA* genes, by RT-PCR and *in silico* analysis. The transcripts of all tested *GmPM* genes showed high accumulation in mid-developmental or pod-dried seeds.

Materials and Methods

Plant materials

Soybean (*Glycine max* L. cv. Shishi) seeds were kindly provided by the Kaohsiung Agricultural Experimental Station, Pintong, Taiwan. Plants were grown to maturity in a greenhouse or field environ-

ment under normal day-length. Germinating cotyledons were sampled around the single-leaf stage. Nodules, roots, stems, leaves or mature flowers were collected during flowering. Pods were harvested at mid-developmental stage, about 35 days after flowering (DAF), and seeds were induced to undergo precocious maturation by air-drying the intact pods (pod-dried) for 4 days as described previously [10]. Soybean seedlings with the first compound leaves fully expanded were cultured for stress treatments in half-strength Johonson's solution [40]. Leaves of control plants or those treated with 100 μM ABA, 50 or 150 mM NaCl, or 2% PEG6000 solution for 1 day, or air-dried for 3 hr at room temperature were harvested. All plant materials were frozen in liquid nitrogen and stored at -80 °C until use.

RNA extraction and analysis

An amount of 500 mg tissue was ground to a fine powder in the presence of liquid nitrogen. Total RNA was extracted by use of TRIzol (ver. 12, Invitrogen) and chloroform, then precipitated with a 0.5 volume of isopropanol and sodium chloride. After washing and cleaning, the RNA was redissolved with double-distilled water. Gene expression analysis involved quantitative RT-PCR. First-strand cDNA synthesis was performed with 0.2 μg total RNA and the Superscript III first-strand cDNA synthesis kit according to the manufacturer's instructions (Invitrogen). Various primers used to detect the endogenous *LEA* genes or elongation factor 1 (*EF-1*) are listed in Table 1. PCR reaction carried out 28 cycles of denaturation for 30 sec at 95 °C; annealing for 45 sec at 54 °C; and extension for 60 sec at 72 °C.

Sequence analysis

Digital Northern analysis of *GmPM* genes involved use of the public soybean expressed sequence tag (EST) database downloaded from NCBI by use of BLAST algorithms [41]. The query sequences were for typical LEA proteins, such

Table 1. Primers used in RT-PCR analysis of *GmPm* genes in soybean.

Clone	Forward	Reverse
GmPM 1	CTGGAACCGCCACGTATTC	TTAAGTACCCCCAGTCCCATACC
GmPM 2	TGCAAGGAGGAAAATGGAAG	GAAGAGAATGTAAACAAGCAAACC
GmPM 3	CTTGGCTTTGCTAGTTGGTGC	GGCTTGTGCATTGGATCTCC
GmPM 4	GAAACCGTTTAAAAACACTATGGG	AACACAGCGTTCACTCAAGAG
GmPM 5	CCATGATTAGCACCTCAACTCC	ACACGTGACCCCAGGTTGTG
GmPM 6	TACTATGGAACCAACACCGC	GCACAACATCTCTGCCTTTATT
GmPM 8	GGTGATGTAGCACAGAATGC	TTCTCAAAGCTCAGCATCCC
GmPM11	GATGAAAGAGCAAGGCAAGGAG	CACATGCTAGTGAAGGGTGTTACAT
GmPM12	TCTGAGGATGATGGGCAAGG	TGCACAAGATACTCAGCCCATG
GmPM16	CAAGTGCTGCCAAGGAACAAG	TGGCCAATGGGACAAGTACC
GmPM17	GTCGCAGTTGTTGGACAAAGC	CCAATGTCCTTTGCCAAGCT
GmPM25	GTGGCGTCACAGTCACAGAAAC	AAGCAGCAGCTGATTGAGCC
GmPM28	ATGGCGAAGAGCAAGGAAGAC	TTTGGGGCTATCCAGGAGCT
GmPM30	AAGCTGGTCAAACTAAGGGCC	CCCTTCACCTTCTCCCCTGT
GmPM39	GTAATGTGGGTGTGGAACAGGAC	CTGTCCCTGGGCTATTCACATAC
EF1*	TGTTGCTGTTAAGGATTTGAAGCG	AACAGTTTGACGCATGTCCCTAAC

*: elongation factor 1 gene

as *GmPM1* (M80666), *2* (M80664), *6* (M94012), *8* (Z22872), *11* (AF004805), *12* (AF004807), *16* (AF004810), *17* (U08108), *25* (AF116754), *28* (AF117724), and *30* (AF117884), and for atypical LEA proteins, such as *GmPM3* (L20806), *4* (U59626), *5* (U59425), and *39* (AF169024).

Results

Tissue-specific expression of GmPM clones

To understand the expression patterns of *GmPM* genes in different tissues, we isolated total RNA from germinating cotyledons, leaves, roots, stems, nodules and flowers from mature plants, as well as mid-developmental or pod-dried seeds, under normal conditions for RT-PCR analysis. All GmPM transcripts could be detected in mid-developmental or pod-dried seeds, with fewer transcripts for GmPM1, 2, 3, 8, and 11 in mid-developmental seeds (Fig. 1). Thus, all tested GmPM genes were LEA-A-type genes [42]. Mid-developmental and pod-dried seeds showed a similar expression level of GmPM4, 5, 6, 25, 28, and 30, which also suggested that the expression of these genes started before the seed water content dropped below 68%-79% [34]. In addition, the transcripts of GmPM4, 6, 17, 25, 30, and 39 were slightly accumulated in mature flowers. Because pollen grains are able to reduce moisture during maturation as seeds develop, these genes might be expressed in the mature pollen grains. Moreover, the expression of most GmPM genes, except for GmPM17 and GmPM25, was barely detectable in vegetative tissues. GmPM4 and 16 were slightly expressed in roots, among all tissues analyzed, whereas the transcripts of GmPM5 and 28 were detected in germinating cotyledons. In addition, the mRNA level of GmPM6 was slightly accumulated in nodules and germinating cotyledons. By contrast, the D-95-type LEA V gene GmPM17 was expressed constitutively in all vegetative tissues, whereas the D-34-type gene GmPM25 was expressed in nodules, roots, or stems.

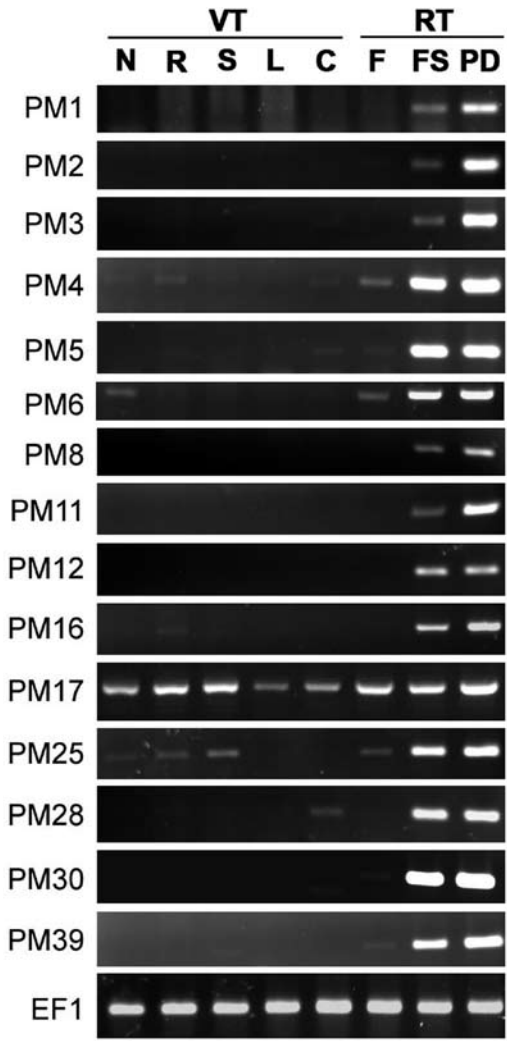

Fig. 1. Tissue-specific expression of *GmPM* (PM) genes in different tissues and developmental stages in soybean. The samples included vegetative tissues (VT), such as nodules (N), roots (R), stem (S), leaves (L), germinating cotyledons (C), as well as reproductive tissues (RT), such as mature flowers (F), 35 DAF seeds (FS), and 4-day pod-dried seeds (PD). Transcripts of soybean elongation factor 1 gene (*EF1*) were used as a control.

ABA or osmotic stress differentially regulated the expression of GmPM genes

The expression patterns of GmPM genes in response to ABA or osmotic stress treatment are illustrated in Figure 2. Although several LEA genes were suggested to be expressed during ABA or osmotic stress, RT-PCR analysis revealed most soybean GmPM genes not expressed under these treatments. Only the transcripts of GmPM5, 6, and 17 were detected under both stresses. The expression level of GmPM5 under ABA or osmotic stress treatment was similar to that of the control, but the level was decreased with NaCl or PEG treatment. Under high salt conditions (150 mM NaCl), the expression of GmPM5 was nearly suppressed. The expression of GmPM17 was decreased with ABA, salt, or PEG treatment. GmPM6 expression was enhanced by high salt or dehydration; however, the level of GmPM6 was barely detected with ABA or PEG treatment and was slightly reduced with weak salt treatment. Thus, the regulation of GmPM6 might be independent of the ABA signaling pathway.

In silico analysis of soybean GmPM clones

The publicly available EST databases are a useful tool for gene functional analysis. At the end of June 2008, the soybean EST database contained 435,347 entries. Hence, we used this database to quantify the expression level of specific genes in various tissues or under various treatments. BLASTN Algorithm was used to search the current soybean EST database, with selected *GmPM* sequences used as queries (see Materials and Methods). In total, we found 537 *GmPM* ESTs from 48 libraries (Table 2). These included 155 ESTs of 68,153 entries for reproductive tissues such as seeds (including cotyledons, seed coats, and seed pods), somatic embryos, and floral tissues, as well as another 314 ESTs of 62,088 entries for biotic or abiotic stress-treated tissue. Only 45 ESTs of 165,485 entries were found for seedlings or vegetative tissues (leaves, roots, or stems), and 64 ESTs of 87,947 entries were from mixed libraries. No library was constructed from

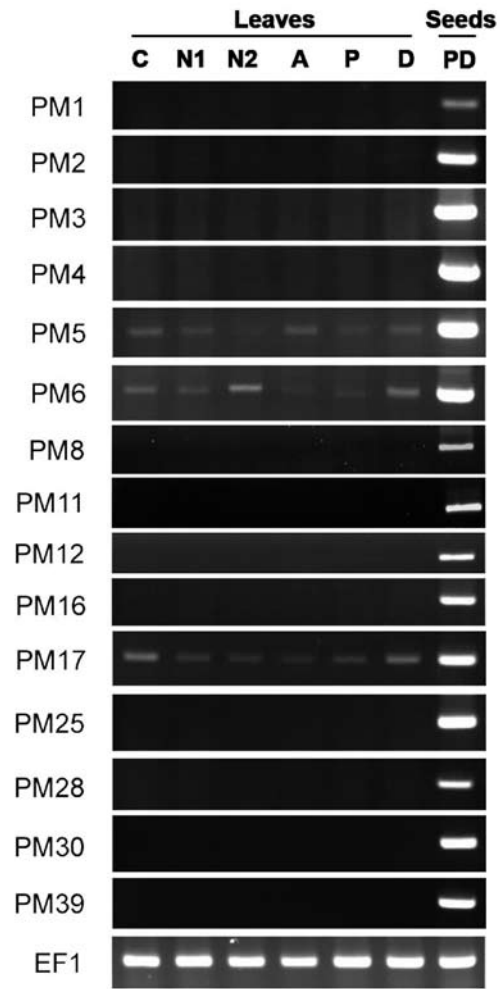

Fig. 2. ABA and osmotic stress differentially regulate the expression of *GmPM* genes. The samples included compound-leaf untreated (C) or treated with 50 mM NaCl (N1), 150 mM NaCl (N2), 100 μM ABA (A), 2% PEG, and drought treatment (D), as well as 4-day pod-dried seeds (PD). Transcripts of soybean *EF1* were used as controls.

Table 2. *In silico* analysis of soybean *GmPM* clones

Library	Treat	PM1	PM2	PM3	PM4	PM5	PM6	PM8	PM11	PM12	PM16	PM17	PM25	PM28	PM30	PM39	Lib. Counts	Total
Gm-c1066	ABS	**8**	0	**2**	0	0	**68**	0	0	0	**2**	0	0	0	**5**	0	**3720**	**85**
Gm-c1068	ABS	**13**	0	**1**	1	0	**178**	0	0	0	**2**	**2**	0	0	**9**	**1**	**5809**	**207**
gmrtDrNS01	ABS	**3**	0	0	0	0	0	0	0	0	**2**	**2**	0	0	**4**	0	**13430**	**11**
Salicylic Acid	ABS	0	0	0	0	0	0	0	0	0	0	**1**	0	0	0	0	**29540**	**1**
H. glycines infection	BS	0	0	0	0	0	**3**	0	0	0	0	0	0	0	**1**	0	**359**	**4**
cyst nematode's infection	BS	0	0	0	0	0	0	0	0	0	0	**1**	0	0	0	0	**444**	**1**
cyst nematode's infection	BS	0	0	0	0	0	0	0	0	0	0	**1**	0	0	0	0	**20**	**1**
P. sojae infection	BS	0	0	0	0	0	0	0	0	0	**1**	**1**	0	0	**2**	0	**3553**	**4**
JCVI-SOY1	M	0	0	**6**	**2**	**3**	**7**	**4**	**1**	**1**	**2**	**3**	0	0	**11**	**1**	**22155**	**41**
JCVI-SOY2	M	0	0	**1**	0	0	**3**	0	0	1	0	**12**	0	0	**1**	0	**21252**	**18**
JCVI-SOY3	M	0	0	0	0	0	0	0	0	0	**1**	**2**	0	0	**2**	0	**17777**	**5**
GmaxSC	R	**1**	**1**	0	0	0	0	0	0	0	0	0	0	0	**3**	0	**2076**	**5**
Gm-c1007	R	0	0	0	0	0	0	0	0	**2**	0	0	0	0	**4**	0	**2396**	**6**
Gm-c1010	R	0	0	0	0	0	0	0	0	**1**	**1**	0	0	0	0	0	**1154**	**2**
Gm-c1011	R	0	0	0	0	0	0	0	0	**2**	0	0	0	0	0	0	**337**	**2**
Gm-c1015	R	0	0	0	0	0	0	0	0	**1**	0	**1**	0	0	0	0	**5827**	**2**
Gm-c1016	R	0	0	0	0	0	0	0	0	0	0	**1**	0	0	0	0	**9146**	**1**
Gm-c1023	R	0	0	**1**	**1**	0	**2**	**1**	0	0	0	**2**	0	0	**1**	0	**3689**	**8**
Gm-c1027	R	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**7656**	**1**
Gm-c1029	R	0	0	0	0	0	0	0	0	0	**1**	0	0	0	**2**	0	**1540**	**3**
Gm-c1036	R	**1**	0	**1**	0	0	**3**	0	0	**9**	**9**	0	**1**	0	**5**	0	**10720**	**29**
Gm-c1051	R	0	0	0	0	0	0	0	0	0	0	**3**	0	0	0	0	**6591**	**3**
Gm-c1055	R	0	0	0	0	**2**	0	0	0	**1**	0	0	0	0	0	0	**3565**	**3**
Gm-c1071	R	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	**4230**	**1**
Gm-c1075	R	0	0	0	0	0	**1**	0	0	**7**	**4**	0	0	0	0	0	**4034**	**12**

Table 2. *In silico* analysis of soybean *GmPM* clones (Continued)

	Library	Treat	PM1	PM2	PM3	PM4	PM5	PM6	PM8	PM11	PM12	PM16	PM17	PM25	PM28	PM30	PM39	Lib. Counts	Total
Gm-c1084		R	0	0	0	0	0	0	0	0	0	0	**1**	0	0	0	0	**4737**	**1**
Gm-c1086		R	0	0	0	0	0	0	0	0	0	0	**1**	0	0	0	0	**1476**	**1**
Gm-r1030		R	**3**	0	0	0	**15**	0	0	0	**6**	**2**	0	0	0	**6**	**2**	**4736**	**34**
Roots		V	0	0	0	0	0	0	0	0	0	0	**1**	0	0	0	0	**997**	**1**
Gm-c1004		V	0	0	0	0	0	0	0	0	0	**1**	**5**	0	0	0	0	**8428**	**6**
Gm-c1009		V	0	0	0	0	0	0	0	0	0	0	**1**	0	0	0	0	**2657**	**1**
Gm-c1018		V	0	0	0	0	0	0	0	0	0	**1**	0	0	0	0	0	**1461**	**1**
Gm-c1028		V	0	0	0	0	0	0	0	0	0	0	**3**	0	0	0	0	**9142**	**3**
Gm-c1039		V	0	0	0	0	0	0	0	0	0	0	**1**	0	0	0	0	**1526**	**1**
Gm-c1040		V	0	0	0	0	**1**	0	0	0	0	0	0	0	0	0	0	**3251**	**1**
Gm-c1048		V	0	0	0	0	0	0	0	0	0	0	**2**	0	0	0	0	**4296**	**2**
Gm-c1049		V	0	0	0	0	0	0	0	0	0	0	**1**	0	0	0	0	**7147**	**1**
Gm-c1052		V	0	0	0	0	0	0	0	0	0	0	**1**	0	0	0	0	**5121**	**1**
Gm-c1053		V	0	0	0	0	0	0	0	0	0	0	**1**	0	0	0	0	**3510**	**1**
Gm-c1054		V	0	0	0	0	0	**2**	0	0	0	0	0	0	0	0	0	**5472**	**2**
Gm-c1062		V	0	0	0	0	0	0	0	0	0	0	**1**	0	0	0	0	**6012**	**1**
Gm-c1063		V	0	0	**1**	0	0	0	0	0	0	**3**	0	0	0	0	0	**4184**	**4**
Gm-c1064		V	0	0	0	0	0	0	0	0	0	**1**	**2**	0	0	0	0	**3091**	**3**
Gm-c1065		V	0	0	0	**1**	0	0	0	0	0	**1**	0	0	0	**1**	0	**8098**	**3**
Gm-c1077		V	0	0	0	0	0	**2**	0	0	0	0	0	0	0	**4**	0	**1782**	**6**
Gm-c1079		V	0	0	0	0	0	0	0	0	0	0	**1**	0	0	0	0	**1331**	**1**
Gm-c1081		V	0	0	0	0	0	0	0	0	0	0	**2**	0	0	**1**	0	**3603**	**3**
Gm-r1021		V	0	0	0	0	0	0	0	0	0	**1**	**2**	0	0	0	0	**3274**	**3**
Total			**34**	**15**	**31**	**15**	**21**	**323**	**12**	**27**	**35**	**42**	**59**	**1**	**1**	**127**	**6**	**278329**	**749**

ABS: abiotic stress; BS: biotic stress; R: reproductive tissues; V: vegetative tissues.

mature or pod-dried seeds in all 146 soybean EST libraries. Hence, the entries of *GmPM* ESTs for reproductive tissues might be low in the present libraries. For example, *GmPM2, 11,* and *25* appeared only once in the EST databases, and no EST was found for *GmPM28*. However, our RT-PCR results showed high accumulation of the transcripts of these genes in mid-developmental or pod-dried seeds. A total of 59 *GmPM17* ESTs appeared in 30 libraries, which suggests that *GmPM17* might be expressed constitutively, which confirmed our RT-PCR results. Salt- and drought-stressed tissues accounted for 292 ESTs of 9,529 entries, including 246 *GmPM6* ESTs. This finding suggested that *GmPM6* genes are highly regulated during salt or drought stress in vegetative tissues, which confirmed our RT-PCR results. Although several *LEA* genes were reported to be involved in the osmotic-response pathway, the EST sequences of 8 tested *GmPM* clones, *GmPM2*, *4*, *5*, *8*, *11*, *12*, *25*, and *28*, were not detected in abiotic stress-treated vegetative tissues.

Discussion

In the present study, we characterized the expression profiles of several soybean *GmPM* genes, including 11 typical and 4 atypical *LEA* genes, by RT-PCR and *in silico* analysis in different tissues or under various stress conditions. All *GmPM* genes were expressed in seeds 35 DAF, which suggests that they are *LEA-A*-type genes. In addition, the expression of all genes was induced in response to seed moisture loss, so they respond to dehydration. The results of RT-PCR or *in silico* analysis suggest that most *GmPM* genes are regulated by spatial effects.

GmPM11 is similar to Em6-type *LEA I* by having only one 20-mer motif [43]. Previous analysis suggested that *LEA I* genes, such as radish *p8B6* [44], wheat *Em* [45], cotton *D-19* [18], barley *B19* [46,47], carrot *Emb-1* [48], Arabidopsis *AtEm1* and *AtEm6* [27,43,49], *Avena fatua AF14* [50], and soybean *Sle1* [51], were seed or embryo specific. Our RT-PCR and *in silico* results confirmed these observations. In addition, Arabidopsis *AtEm1* transcripts were detected about 13-14 days after pollination and *AtEm6* transcripts 3 days later, which indicates that the *AtEm1* gene belongs to *LEA-A*-type genes, whereas *AtEm6* is an *LEA* type, and *GmPM11* performs a *LEA-A*-type temporal regulation. Thus, the number of 20-mer motifs might not be related to the regulation of the *Em* gene.

Two soybean *LEA II* genes, *GmPM6* (Y_2K type) and *GmPM12* (Y_3SK type), were shown to be *LEA-A* type in mid-developmental seeds. However, in vegetative tissues, these two genes had different expression patterns. Both RT-PCR and *in silico* analysis indicated that *GmPM6* was expressed in vegetative tissues and mature flowers and the expression was induced in response to high salt or dehydration. By contrast, *GmPM12* is seed or embryo specific, and transcripts did not accumulate under ABA or stress treatment. Previous studies had suggested that dehydrins in the same group might or might not give similar organ-, development-, or stress-specific expression patterns. For instance, 5 Arabidopsis *dehydrin* genes -- *Cor47* (SK_3 type), *Lti29* (SK_2 type), *Erd14* (SK_2 type), *Lti30* (K_6 type), and *Rab18* (Y_2SK_2 type) -- show differential expression patterns in response to low temperature, salinity, and ABA treatment in root, stem and leaf tissues. The accumulation of *Erd14* and *Lti30* mRNA is constitutive, but *Lti30* is induced in response to low temperature or salinity in tested tissues. *Cor47* and *Lti29* are mainly expressed in roots and stems and slightly expressed in leaves under low temperature. *Rab18* mRNA accumulates only in ABA- and salinity-treated tissues. As is common, the accumulation of *dehydrin* genes is higher in roots and stems than in leaves [22,52,53]. Four members of the rice

Rab family (*Rab21*, *rab16B* to *D*), which are arrayed in tandem on chromosome 11 and have similar sequences (YSK_2 type), are expressed in ABA-treated seedlings, and only *Rab16D* is undetected in mature seeds [13,54]. Thus, the expression patterns of *dehydrins* reveal complex regulation pathways in seeds or under environmental stresses.

Three soybean *LEA III* genes, with their deduced protein sequences containing different numbers and consensus sequences of 11-mer repeats, were used to examine expression profiles. The deduced protein sequence of *GmPM2* genes contains 19 11-mer repeats and one additional 36-mer motif. RT-PCR and *in silico* results suggested that *GmPM2* is seed or embryo specific, although the expression of its wheat ortholog *MA1949* is induced by water stress in seedlings [55]. The deduced protein sequence of *GmPM8* shows 32 11-mer repeats and a signal peptide in the N terminal. Our previous northern blot analysis revealed that the transcripts of the *GmPM8* homolog *GmPM10* accumulate in callus or seedlings under drought or chilling treatment, as well as in seeds [15]. In the present study, however, the transcripts of *GmPM8* were detected only in mid-developmental seeds. *GmPM30* is related to D-7-type *LEA* genes**,** with five 11-mer repeats. The transcripts of *GmPM30* were detected in mid-developmental seeds. *In silico* results suggested that *GmPM30* might be expressed under abiotic stress. Considering the low level of EST entries for this gene -- 5, 9, and 4 ESTs of 3,720; 5,809; and 13,430 entries, respectively -- *GmPM30* transcripts possibly were not detected by RT-PCR. Animal orthologs of *GmPM30* were suggested to be induced in response to anhydrobiotic stress [e.g. 56-59]. In conclusion, although numbers of *LEA III* genes are responsive to various environmental stresses, the tested soybean *LEA III* genes are mainly expressed in seeds 35 DAF.

We used three different *LEA IV* genes -- *GmPM1*, *GmPM16* and *GmPM28* -- according to their conserved N-terminal domains to examine their expression profiles. RT-PCR results showed these genes highly expressed in pod-dried seeds and their expression not induced in response to exogenous ABA or osmotic stress treatment, although *in silico* analysis showed small amounts of *GmPM1* or *GmPM16* ESTs. *LEA IV* genes have been isolated from several plant species, including cotton, cultivar or wild-type tomato, sunflower, soybean, *Lemna gibba*, Arabidopsis, *Phaseolus vulgaris*, white spruce, birch, and *C. plantagineum*, as well as genome sequences from various plant species [see reviewed in 1,3,4]. Unlike the protein sequences of *LEA I* to *III* genes, the deduced protein sequences of *LEA IV* genes lack consensus motifs or sequences. In general, the N-terminal domains contain several conserved small and charged amino acid residues, whereas the variable C-terminal regions include a large number of non-polar amino acid residues. In most studies, *LEA IV* genes were expressed in both mature seeds and stress-treated vegetative tissues [e.g. 16,25,60-62]. Arabidopsis *PAP51*, however, is seed specific [49].

As typically hydrophobic *LEA* genes, D-95-type *GmPM17* and D-34-type *GmPM25* displayed expression patterns different from those of typically hydrophilic *LEA* genes. The transcripts of *GmPM17* were detected in all tested tissues, although the accumulation in leaves or germinating cotyledons was lower than that in other tissues. *In silico* analysis also confirmed the observation. However, the expression of *GmPM17* was decreased under exogenous ABA or osmotic stress treatment in vegetative tissue. The transcripts of *GmPM25* accumulate in roots, stems, flowers, or seeds. No *GmPM25* transcripts were detected in leaves under normal or stress-treated conditions. We found only one *GmPM25* hit in the whole EST database (435,347 entries).

This finding might be explained by random selection for each cDNA library or the tissue to construct the library. Previous analysis showed several *LEA V* genes, such as cotton *D-34* and *D-95*, carrot *ECP31*, or Arabidopsis *PAP140* and *AtRAB28*, expressed in seeds or induced in young embryos by dehydration [18,49,63,64]. The expression of other *LEA V* genes, such as *C. Plantagineum pcC27-45* and *pcC16-81*, maize *RAB28*, tomato *ER5*, and hot pepper *CaLEA6*, was induced or enhanced by salt or dehydration in vegetative tissues [24,26,65-68]. Soybean *GmPM17* and *GmPM25* and Arabidopsis *Atecp31* [64] are examples of *LEA V* genes expressed both in seeds and vegetative tissues.

Similar to the protein sequences of *LEA V* genes, the deduced sequences of 4 tested atypical *LEA* genes are hydrophobic. The *GmPM3* homolog, wheat *AWPM-19*, was suggested to be involved in freezing tolerance [69]. The deduced proteins of GmPM3 and AWPM-19 contain a conserved domain that may act as transmembrane domains. In our study, the transcripts of *GmPM3* were detected only in mid-developmental or pod-dried seeds and were absent under exogenous ABA or osmotic stress treatment. GmPM4 protein and its homolog pea SBP65 are biotinylated proteins and mainly accumulate in mature seeds [37,70]. Our RT-PCR analysis revealed the transcripts of *GmPM4* in nodules, roots, and mature flowers under normal conditions and not in seedlings under exogenous ABA or osmotic stress treatment. The product of *GmPM5* was annotated as a 7S seed globulin precursor, with its transcripts slightly accumulating in germinating cotyledons or leaves under normal conditions and detected in seedlings under exogenous ABA or osmotic stress treatment. Severe salt stress might suppress the expression of *GmPM5*. RT-PCR and *in silico* results revealed *GmPM39* expressed in seeds or mature flowers under normal conditions and not ABA- or stress-treated conditions. However, the *GmPM39* homolog spinach *CAP160* showed the response under chilling or severe water stress [71]. Hence, these two genes might be involved in diverse regulation pathways.

Acknowledgements

We express our deep thanks to Ms. Laura Heraty for editing services. This research was supported by National Science Council and Academia Sinica, Taiwan, to YIC Hsing.

References

[1] Dure L, Crouch M, Harada J, Ho THD, Mundy J, Quatrano R, Thomas T, Sung ZR: Common amino acid sequence domains among the LEA proteins of higher plants. Plant Molecular Biology 1989; 12, 475-486.

[2] Ingram J, Bartels D: The molecular basis of dehydration tolerance in plants. Annual Review of Plant Physiology and Plant Molecular Biology 1996; 47, 377-403.

[3] Cuming AC: LEA proteins. In Seed Proteins Edited by Shewry PR, Casey R. Kluwer Academic Publishers, Dordrecht, Netherlands. 1999; 753-780.

[4] Shih MD, Hoekstra, FA, Hsing YIC: Late embryogenesis abundant proteins. Advance in Botanical Research 2008; *in press*

[5] Tunnacliffe A Wise MJ: The continuing conundrum of the LEA proteins. Naturwissenschaften (2007); 94, 791-812.

[6] Dure III L: Structural motifs in Lea proteins. In Plant responses to cellular dehydration during environmental stress Edited by Close TJ, and Bray EA. American Society of Plant Physiologyogists Series. USA. 1993; 91-103

[7] Wise MJ: LEAping to conclusions: A computational reanalysis of late embryogenesis abundant proteins and their possible roles. BMC Bioinformatics 2003; 4, 52-70.

[8] Wolkers WF, McCready S, Brandt WF, Lindsey GG, Hoekstra FA: Isolation and characterization of a D-7 LEA protein from pollen that stabilizes glasses *in vitro*. Biochimica et Biophysica Acta 2001; 1544, 196-206.

[9] Goyal K, Tisi L, Basran A, Browne J, Burnell A, Zurdo J, Tunnacliffe A:

Transition from natively unfolded to folded state induced by desiccation in an anhydrobiotic nematode protein. Journal of Biological Chemistry 2003; 278, 12977-12984.

[10] Shih MD, Lin SC, Hsieh JS, Tsou CH, Chow TY, Lin TP, Hsing YIC: Gene cloning and characterization of a soybean (*Glycine max* L.) LEA protein, GmPM16. Plant Molecular Biology 2004; 56, 689-703.

[11] Boudet J, Buitink J, Hoekstra FA, Rogniaux H, Larre C, Satour P, Leprince O: Comparative analysis of the heat stable proteome of radicles of *Medicago truncatula* seeds during germination identifies late embryogenesis abundant proteins associated with desiccation tolerance. Plant Physiology 2006;140, 1418-1436.

[12] Singh S, Cornilescu CC, Tyler RC, Cornilescu G, Tonelli M, Lee MS, Markley JL: Solution structure of a late embryogenesis abundant protein (LEA14) from *Arabidopsis thaliana*, a cellular stress-related protein. Protein Science 2005; 14, 2601-2609.

[13] Mundy J, Chua NH: Abscisic acid and water-stress induce the expression of a novel rice gene. EMBO Journal 1988; 7, 2279-2286.

[14] Hong B, Barg R, Ho THD: Developmental and organ-specific expression of an ABA- and stress-induced protein in barley. Plant Molecular Biology 1992; 18, 663-674.

[15] Hsing YIC, Chen ZY, Shih MD, Hsieh JS, Chow TY: Unusual sequences of group 3 LEA mRNA inducible by maturation or drying in soybean seeds. Plant Molecular Biology 1995; 29, 863-868.

[16] Cohen A, Bray EA: Nucleotide sequence of an ABA-induced tomato gene that is expressed in wilted vegetative organs and developing seeds. Plant Molecular Biology 1992; 18, 411-413.

[17] Morris PC, Kumar A, Bowles DJ, Cuming AC: Osmotic stress and abscisic acid induce expression of the wheat *Em* genes. European Journal of Biochemistry 1990; 190, 625-630.

[18] Hughes DW, Galau GA: Developmental and environmental induction of *Lea* and *LeaA* mRNAs and the postabscission program during embryo culture. Plant Cell 1991; 3, 605-618.

[19] Choi JH, Liu LS, Borkird C, Sung ZR: Cloning of genes developmentally regulated during plant embryogenesis. Proceedings of the National Academy of Sciences of the United States of America 1987; 84, 1906-1910.

[20] Harada JJ, DeLisle AJ, Baden CS, Crouch ML: Unusual sequence of an abscisic acid-inducible mRNA which accumulates late in *Brassica napus* seed development. Plant Molecular Biology 1989; 12, 395-401.

[21] Gilmour SJ, Artus NN, Thomashow MF: cDNA sequence analysis and expression of two cold-regulated genes of Arabidopsis thaliana. Plant Molecular Biology 1992; 18, 13-21.

[22] Nylander M, Svensson J, Palva ET, Welin BV: Stress-induced accumulation and tissue-specific localization of dehydrins in *Arabidopsis thaliana*. Plant Molecular Biology 2001; 45, 263-279.

[23] Neven LG, Haskell DW, Hofig A, Li QB, Guy CL: Characterization of a spinach gene responsive to low temperature and water stress. Plant Molecular Biology 1993; 21, 291-305.

[24] Piatkowski D, Schneider H, Salamini F, Bartels D: Characterization of five abscisic acid-responsive cDNA clones isolated from the desiccation-tolerance plant *Craterostigma plantagineum* and their relationship to other water-stress genes. Plant Physiology 1990; 94, 1682-1688.

[25] Ditzer A., Kirch HH, Nair A, and Bartels D: Molecular characterization of two alanine-rich *Lea* genes abundantly expressed in the resurrection plant *C. plantagineum* in response to osmotic stress and ABA. Journal of Plant Physiology 2001; 158, 623-633.

[26] Smith-Espinoza CJ, Richter A, Salamini F, Bartels D: Dissecting the response to dehydration and salt (NaCl) in the resurrection plant *Craterostigma plantagineum*. Plant, Cell & Environment 2003; 26, 1307-1315.

[27] Bies N, Aspart L, Carles C, Gallois P, Delseny M: Accumulation and degradation of Em proteins in *Arabidopsis thaliana*: Evidence for post-transcriptional controls. Journal of Experimental Botany 1998; 49, 1925-1933.

[28] Vicient CM, Gruber V, Delseny M: The Arabidopsis *AtEm1* promoter is active in *Brassica napus* L. and is temporally and spatially regulated. Journal of Experimental Botany 2001; 52, 1587-1591.

[29] Xu D, Duan X, Wang B, Hong B, Ho THD, Wu R: Expression of a late embryogenesis

abundant protein gene, *HVA1*, from barley confers tolerance to water deficit and salt stress in transgenic rice. Plant Physiology 1996; 110, 249-257.
[30] Sivamani E, Bahieldinl A, Wraith JM, Al Niemi T, Dyer WE, Ho THD, Qu R: Improved biomass productivity and water use efficiency under water deficit conditions in transgenic wheat constitutively expressing the barley *HVA1* gene. Protein Science 2000; 155, 1-9.
[31] Maqbool B, Zhong H, El Maghraby Y, Ahmad A, Chai B, Wang W, Sabzikar R, Sticklen B: Competence of oat (*Avena sativa* L.) shoot apical meristems for integrative transformation, inherited expression and osmotic tolerance of transgenic lines containing *hva1*. Theoretical and Applied Genetics 2002; 105, 201-208.
[32] Babu RC, Zhang J, Blum A, Ho THD, Wu R, Nguyen HT: *HVA1*, a *LEA* gene from barley confers dehydration tolerance in transgenic rice (*Oryza sativa* L.) via cell membrane protection. Protein Science 2004; 166, 855-862
[33] Armstrong C, Black M, Chapman JM, Norman HA, Angold R: The induction of sensitivity to gibberellin in aleurone tissue of developing wheat grains. Planta 1982; 154: 573-577.
[34] Rosenberg LA, Rinne RW: Moisture loss as a prerequisite for seedling growth in soybean seeds (*Glycine max* L. Merr.). Journal of Experimental Botany 1986; 37, 1663-1674.
[35] Rosenberg LA, Rinne RW: Protein synthesis during rehydration, germination and seedling growth of naturally and precociously matured soybean seeds (*Glycine max*). Annals of Botany 1989; 64, 77-86
[36] Hsing YIC, Wu S: Cloning and characterization of cDNA clones encoding soybean seed maturation polypeptides. Botanical Bulletin of Academia Sinica 1992; 33, 191-199
[37] Hsing YIC, Tsou CH, Hsu TF, Chen ZY, Hsieh KL, Hsieh JS, Chow TY: Tissue- and stage-specific expression of a soybean (*Glycine max* L.) seed-maturation, biotinylated protein. Plant Molecular Biology 1998; 38: 491-490
[38] Lee PF, Hsing YIC, Chow TY: Promoter activity of a soybean gene encoding a seed maturation protein, *GmPM9*. Botanical Bulletin of Academia Sinica 2000; 41, 175-182
[39] Shih MD, Hsieh JS, Hsing YIC: Regulation of the Soybean *GmPM9* Promoter in Callus Tissue. Journal of Genetics and Molecular Biology 2001; 12, 125-137.
[40] Hewitt EJ: Sand and water culture method used in study of plant nutrition. Technical Communication No.22 (Revised), Commonwealth Bureau of Horticulture and Plantation Crops, East Mailing Maidstone, Kent, England. 1966.
[41] Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research 1997; 25, 3389-3402.
[42] Hughes DW, Galau GA: Temporally modular gene expression during cotyledon development. Genes & Development 1989; 3, 358-369.
[43] Gaubier P, Raynal M, Hull G, Huestis GM, Grellet F, Arenas C, Pages M, Delseny M: Two different *Em*-like genes are expressed in *Arabidopsis thaliana* seeds during maturation. Molecular & General Genetics 1993; 238, 409-418.
[44] Raynal M, Depigny D, Cooke R, Delseny M: Characterization of a radish nuclear gene expressed during late seed maturation. Plant Physiology 1989; 91, 829-836.
[45] Morris PC, Kumar A, Bowles DJ, Cuming AC: Osmotic stress and abscisic acid induce expression of the wheat *Em* genes. European Journal of Biochemistry 1990; 190, 625-630
[46] Espelund M, Saeboe-Larssen S, Hughes DW, Galau GA, Larsen F, Jakobsen KS: Late embryogenesis-abundant genes encoding proteins with different numbers of hydrophilic repeats are regulated differentially by abscisic acid and osmotic stress. Plant Journal 1992; 2, 241-252
[47] Espelund M, De Bebout JA, Outlaw Jr. WH, Jakobsen KS: Environmental and hormonal regulation of barley late-embryogenesis-abundant (*Lea*) mRNAs is via different signal transduction pathways. Plant, Cell & Environment 1995; 18, 943-949.
[48] Wurtele ES, Wang H, Durgerian S, Nikolau BJ, Ulrich TH: Characterization of a gene that is expressed early in somatic embryogenesis of *Daucus carota*. Plant Physiology 1993; 102, 303-312.
[49] Parcy F, Valon C, Raynal M, Gaubier-Comella P, Delseny M, Giraudat J:

Regulation of gene expression programs during Arabidopsis seed development: roles of the *ABI3* locus and of endogenous abscisic acid. Plant Cell 1994; 6, 1567-1582.

[50] Li B, Foley ME: Cloning and characterization of differentially expressed genes in imbibed dormant and afterripened *Avena fatua* embryos. Plant Molecular Biology 1995; 29, 823-831.

[51] Calvo ES, Wurtle ES, Shoemaker RC: Cloning, mapping, and analyses of expression of the *Em*-like gene family in soybean [*Glycine max* (L) Merr]. Theoretical and Applied Genetics 1997; 94, 957-967

[52] Kiyosue T, Yamaguchi-Shinozaki K, Shinozaki K: Characterization of two cDNAs (*ERD10* and *ERD14*) corresponding to genes that respond rapidly to dehydration stress in *Arabidopsis thaliana*. Plant & Cell Physiology 1994; 35, 225-231

[53] Welin BV, Olson A, Nylander M, Palva ET: Characterization and differential expression of *dhn*/*lea*/*rab*-like genes during cold acclimation and drought stress in *Arabidopsis thaliana*. Plant Molecular Biology 1994; 26, 131-144.

[54] Yamaguchi-Shinozaki K, Mundy J, Chua NH: Four tightly linked rab genes are differentially expressed in rice. Plant Molecular Biology 1990; 14, 29-39.

[55] Curry J, Walker-Simmons MK: Unusual sequence of group 3 *LEA* (II) mRNA inducible by dehydration stress in wheat. Plant Molecular Biology 1993; 21, 907-912.

[56] Browne J, Tunnacliffe A, Burnell A: Anhydrobiosis: plant desiccation gene found in a nematode. Nature 2002; 416, 38.

[57] Tunnacliffe A, Lapinski J, Mcgee B: A putative LEA protein, but no trehalose, is present in anhydrobiotic bdelloid rotifers. Hydrobiologia 2005; 546, 315-321.

[58] Kikawada T, Nakahara Y, Kanamori Y, Iwata K, Watanabe M, McGee B, Tunnacliffe A, Okuda T: Dehydration-induced expression of LEA proteins in an anhydrobiotic chironomid. Biochemical and Biophysical Research Communications 2006; 348, 56-61.

[59] Hand SC, Jones D, Menze MA, Witt TL: Life without water: expression of plant *LEA* genes by an anhydrobiotic arthropod. Journal of Experimental Zoology - Part A: Ecological Genetics and Physiology 2006; 307, 62-66.

[60] Kahn TL, Fender SE, Bray EA, O'Connell MA: Characterization of expression of drought- and abscisic acid-regulated tomato genes in the drought-resistant species *Lycopersicon pennellii*. Plant Physiology 1993; 103, 597-605

[61] Colmenero-Flores JM, Campos F, Garciarrubio A, Covarrubias AA:. Characterization of *Phaseolus vulgaris* cDNA clones responsive to water deficit: identification of a novel late embryogenesis abundant-like protein. Plant Molecular Biology 1997; 35, 393-405.

[62] Colmenero-Flores JM, Moreno LP, Smith CE, Covarrubias AA: *Pvlea-18*, a member of a new late-embryogenesis-abundant protein family that accumulates during water stress and in the growing regions of well-irrigated bean seedlings. Plant Physiology 1999; 120, 93-104.

[63] Kiyosue T, Yamaguchi-Shinozaki K, Shinozaki K, Higashi K, Satoh S, Kamada H, Harada H: Isolation and characterization of a cDNA that encodes *ECP31*, an embryogenic-cell protein from carrot. Plant Molecular Biology 1992; 19, 239-249

[64] Arenas-Mena C, Raynal M, Borrell A, Varoquaux F, Cutanda MC, Stacy RA, Pages M, Delseny M, Culianez-Macia FA: Expression and cellular localization of *Atrab28* during Arabidopsis embryogenesis. Plant Molecular Biology 1999; 40, 355-363

[65] Michel D, Furini A, Salamini F, Bartels D: Structure and regulation of an ABA- and desiccation-responsive gene from the resurrection plant *Craterostigma plantagineum*. Plant Molecular Biology 1994; 24, 549-560.

[66] Niogret MF, Culianez-Macia FA, Goday A, Mar AM, Pages M: Expression and cellular localization of *rab28* mRNA and Rab28 protein during maize embryogenesis. Plant Journal 1996; 9, 549-557

[67] Zegzouti H, Jones B, Marty C, Lelievre JM, Latche A, Pech JC, Bouzayen M: ER5, a tomato cDNA encoding an ethylene-responsive LEA-like protein: characterization and expression in response to drought, ABA and wounding. Plant Molecular Biology 1997; 35, 847-854.

[68] Kim HS, Lee JH, Kim JJ, Kim CH, Jun SS, Hong YN: Molecular and functional characterization of *CaLEA6*, the gene for a hydrophobic LEA protein from *Capsicum annuum*. Gene 2005; 344, 115-123.

[69] Koike M, Takezawa D, Arakawa K, Yoshida S: Accumulation of 19-kDa plasma mem-

brane polypeptide during induction of freezing tolerance in wheat suspension-cultured cells by abscisic acid. Plant & Cell Physiology 1997; 38, 707-716.

[70] Duval M, Job C, Alban C, Douce R, Job D: Developmental patterns of free and protein-bound biotin during maturation and germination of seeds of *Pisum sativum*: characterization of a novel seed-specific biotinylated protein. Biochemistry Journal 1994; 299, 141-150.

[71] Kaye C, Neven L, Hofig A, Li QB, Haskell D, Guy C: Characterization of a gene for spinach *CAP160* and expression of two spinach cold-acclimation proteins in tobacco. Plant Physiology 1998; 116, 1367-1377.