Fruit Ripening-Related Expression of a Gene Encoding Group 5 Late **Embryogenesis Abundant Protein in Citrus**

Oh-Sung Jeon¹, Chan-Shick Kim¹, Sam-Pin Lee², Sung Ku Kang³, Chang-Myung Kim⁴, Bong-Gyoon Kang⁵, Yoonkang Hur^{6*}, and In-Jung Kim^{1, 7}

¹Faculty of Biotechnology, College of Applied Life Sciences, Cheju National University, Jeju 690-756, Korea ²Department of Food Science and Technology, Keimyung University, Daegu 704-701, Korea ³Korea National Agricultural College, Hwasung 445-890, Korea ⁴Citrus Division, National Institute of Subtropical Agriculture, RDA, Jeju 690-150, Korea

⁵The Center for the Scientific Technology of Citrus and Flowers, Cheju National University, Jeju 690-756, Korea

⁶Department of Biology, School of Bioscience and Biotechnology, Chungnam National University, Daejeon 305-764, Korea 7Applied Radiological Science Research Institute, Cheju National University, Jeju 690-756, Korea

A cDNA and genomic clone (CuLEA5) encoding a group 5 late embryogenesis abundant protein (Lea5) was isolated from citrus fruit cDNA and genomic libraries. Sequence analysis indicated that the clone contains an open reading frame of 97 amino acids, and that the genomic structure is composed of two exons and one intron. A comparison of its amino sequence with other plant proteins showed that Lea5 proteins can be classified into two types – gymnosperm and angiosperm – based on a P-segment sequence designated by this study. Examination of its expression patterns indicated that CuLEA5 has important roles during the development or ripening of seedless fruits and leaves in Citrus. The 5'-flanking region of the genomic DNA contains a number of putative hormonal- and stress-responsive elements. This is the first report that describes the expression of Lea5 during fruit ripening, as well as the sequence characteristics of its promoter region.

Keywords: cDNA, expression, fruit ripening, genomic DNA, group 5 late embryogenesis abundant protein

Late embryogenesis abundant (LEA) proteins accumulate both during the late stages of seed development and in other plant tissues. This induction can be prompted by exposure to exogenous abscisic acid (ABA), desiccation, osmosis, and low temperatures (Wang et al., 2003; Bernacchia and Furini, 2004; Shao et al., 2005). The exact roles of LEA proteins are mostly unknown. Nevertheless, functioning predictions are facilitated because of their considerable synthesis during the late period of embryogenesis, the induction by stress, and various structural characteristics. It has been suggested that LEAs protect desiccating tissues from the effects of water loss by binding water, sequestering ions, guarding proteins and membranes, and re-naturizing unfolded proteins (Shao et al., 2005).

LEA proteins were first studied in developing cotton seeds, where they were categorized into five or six groups based on their amino acid sequence, expression, and structural features (Wise, 2003; Shao et al., 2005). The major categories are Groups 1, 2, and 3, with Group 2 members -'dehydrin' and the 'Lea D11 family' - being the most frequently studied (Close, 1997; Ha et al., 2006). Most LEA proteins are composed of hydrophilic and biased amino acids, and have highly conserved amino acid motifs. Significantly increased osmotic stress tolerance has resulted from the overexpression of Groups 2 and 3 LEA proteins in several plants, including Arabidopsis (Artus et al., 1996; Ndong et al., 2002), radish (Park et al., 2005), rice (Xu et al., 1996), tobacco (Hara et al., 2003), and wheat (Sivamani et al., 2000).

Group 5 LEA proteins (Lea5) have been described from

cotton (Galau et al., 1993) and soybean (Burns et al., 1996) upon seed maturation, from potato upon tuberization (Jackson et al., 1997), from mung bean treated with auxin (Yamamoto, 1994), from the developing roots of pineapple (Neuteboom et al., 2002), and from hot pepper (Hwang et al., 2005) and Citrus (Naot et al., 1995) under external stresses. Their sequences commonly display unremarkable hydropathic profiles that differ from other LEA protein groups. These Lea5 proteins contain hydrophobic domains at the N-terminal half and the hydrophilic C-terminal half.

Numerous physiological changes, including compositional, structural, and metabolic alterations, occur in Citrus during fruit ripening (Baldwin, 1993; Giovannoni, 2004). Many of these can be explained by fluctuations in gene expression. Several genes involved in carotenoid biosynthesis and stress responses may accumulate at that maturation stage (Kim et al., 2001a, b). Except for seed formation, little information is available concerning the expression and regulation of a LEA gene during fruit development. Several coldand salt-responsive genes related to Group 2 (dehydrin) (Hara et al., 1999; Porat et al., 2002, 2004) and Group 5 LEA proteins (Naot et al., 1995) have now been cloned from several Citrus species. Their expressions are increased by drought, heat, salt, and cold stresses. However, none have been investigated for their activation during fruit ripening. For example, Satsuma mandarin is a male-sterile cultivar that, unlike other Citrus relatives (e.g., orange and grapefruit), is able to set parthenocarpic (seedless) fruit, producing only fleshy parts and peels. Thus, studies on the relationship between fruit ripening and expression of the Lea5 gene in Satsuma mandarin can lead to an expansion of our understanding of the functioning of Lea5 proteins during normal development outside of seed formation. Here, we isolated Culea5 cDNA and genomic DNA as fruit-dominant genes,

^{*}Corresponding author; fax +82-64-756-3351, +82-42-822-9690 e-mail ijkim@čheju.ac.kr, ykhur@cnu.ac.kr

and characterized them as fruit ripening-related genes in *Citrus*.

MATERIALS AND METHODS

Plant Materials

Satsuma mandarin (*Citrus unshiu* Marc. cv. Miyagawa Wase), hereafter referred to as *Citrus*, was cultivated at the Citrus Experiment Station in Jeju (Korea). The fruits and leaves of greenhouse-grown plants were sampled. Fruits were harvested from the immature stage (July) to the mature stage (October) for analysis of expression patterns at the RNA level.

Construction of a cDNA Library and Differential Screening

Total RNA was extracted from full, yellow fruits and leaves, using the hot phenol RNA isolation procedure (Verwoerd et al., 1989). *Citrus* cDNA libraries were constructed as described by Kim et al. (2001a). Duplicate filters from the *Citrus* fruit cDNA library were first hybridized to cDNA probes synthesized from 1 mg poly(A) RNA, either from the leaves or the mature fruits. The clones were isolated based on their ability to perform preferential hybridization to the probe of ripe fruits. The cDNA probes were then labeled with [α -³²P]dCTP, using the Prime-a-Gene labeling system (Promega, USA).

DNA Sequence Analyses

The isolated cDNA and genomic DNA were sequenced on both strands with T3, T7, and sequence-specific primers. Multiple sequence alignments were performed with the ClustalX (1.8) computer program (Thompson et al., 1997). Aligned sequences from the deduced amino acid sequences were then used to build a phylogenetic tree according to the TreeviewX 1.6.6 program.

Northern and Southern Blot Analyses

For northern blot analysis, total RNA was isolated from fruits at seven developmental stages (53, 67, 95, 109, 123, 137, and 153 days after fertilization; DAF), as well as from oneyear-old leaves. For our Southern blot analysis, genomic DNA was obtained from young leaves, according to the method of Dellaporta et al. (1983). The genomic DNA was digested with *Bam*HI, *XhoI*, and *Eco*RI. Hybridization and washing of filters were done as described by Kim et al. (2001a). The probe was *CuLEA5* cDNA labeled with [α -³²P] dCTP.

RT-PCR

DNase I-treated total RNA (50 ng) from old (one-year-old) and young (3- to 6-month-old) leaves was used as template in RT-PCR (AccuPower RT/PCR PreMix; Bioneer, Korea). The reactions were performed with gene-specific primers with 25 cycles of 30 s 94°C, 30 s 60°C, and 30 s 72°C, that was first preceded by 5 min 94°C, and followed by 5 min at 72°C. PCR reactions were normalized by including a primer pair to amplify a 286-bp fragment of 18S rRNA in the same tube. The specifically designed primers for PCR included:

LEA5-F, 5'-AGCCGCTAGCTTTTGCCTAT-3' and LEA5-R, 5'-CCGCTAACAGCAAAGGATCT-3' for *CuLEA5*; and rRNA-F, 5'-GAACAACTGCGAAAGCATTTGC-3' and rRNA-R, 5'-CCTGGTAAGTTTCCCCGTGTTG-3' for 18S rRNA.

Construction and Screening of a Genomic Library

Genomic DNA was partially digested with *Sau*3AI, ligated with the ZAP Express *Bam*HI-Predigested Vector (Stratagene, USA), and packaged into GigapackIII gold cloning kits (Stratagene), according to the manufacturer's instructions. The library was then screened with a radiolabeled *CuLEA5* cDNA probe, labeled with $[\alpha$ -³²P]dCTP.

RESULTS AND DISCUSSION

We have isolated a *CuLEA5* cDNA clone from a *Citrus* fruit cDNA library via differential screening on the basis of its increased expression in fruits as compared to leaves (data not shown). This clone (GenBank accession number DQ424890) is 583 bp long. The *CuLEA5* cDNA contains a 294-b open reading frame, a 127-b 5'-UTR (untranslated region), and a 162-b 3'-UTR. The coding region is composed of 97 amino acids deriving about 10.6 kDa of polypeptide (Fig. 1A). Moreover, a putative polyadenylation signal is found at positions 558-563 (AATAGT).

Thirteen nucleotide differences exist between the cDNA sequences of the sweet orange *C-Lea5* and Satsuma mandarin *CuLEA5*, leading to two amino acid substitutions between the proteins (Fig. 1A). In addition, most of the different nucleotides are found at the 5'- and 3'-UTRs. This indicates that Satsuma mandarin *CuLEA5* is a homologue of orange *C-Lea5*.

We compared the deduced amino acid sequences between CuLEA5 and other plant Lea5 proteins (Fig. 1). The protein sequences show a high degree of conservation, especially at the amino terminus and in the C-terminal half of the protein (Fig. 1A, boxed by dotted line). The N-terminal region of the protein contains the characteristics of a signal peptide, as predicted by the SignalP 3.0 server (www.cbs.dtu.ck/services/SignalP). This signal peptide region is necessary for facilitating the targeting of the endoplasmic reticulum and cell secretion. The C-terminal half contains a 12-amino acid stretch, designated as a P-segment, which has the consensus W(A/L/V)PDP(V/I)TG(Y/H)YRP in angiosperm plants and WMRDP(A/T)TG(N/D)WIP in gymnosperms. These can be referred to as angiosperm-type and gymnosperm-type P-segments, respectively. A similar classification has also been found in the case of the K-segment of the dehydrin, Group 2 Lea proteins (Close, 1997). These sequence characteristics suggest that Lea5 proteins have a similar function, although this is yet to be elucidated.

The *CuLEA5* sequence has a high degree of similarity with those of various angiosperms (30 to 57%), but not with gymnosperm species (26 to 32%). To date, gymnosperm Lea5 has been isolated only from spruce. CuLEA5 generally possesses stronger similarities with dicotyledonous plants (40 to 57%) than with monocots, including barley (30%) and rice (32%). Our phylogenetic dendrogram indicated that *Lea5* can be categorized into two branches of angiosperm and



Figure 1. Comparison of deduced amino acid sequences (**A**) and phylogenic relationship (**B**) of *Lea5* genes among plant species. Gene identities are described by plant names and GenBank accession numbers. Predicted amino acid sequences were aligned using ClustalX (version 1.64b). Gaps are marked with dashes; conserved amino acid residues, with asterisks. Conserved regions at N- and C-terminal half are boxed and shaded. P-segment is located within conserved C-terminal half box. Arrow heads indicate conserved amino acid in angiosperm plants but not in gymnosperm plants. Tree was generated by ClustalX and TreeView (version 1.6.1).



Figure 2. Genomic (A), northern blot (B), and RT-PCR (C) analyses of *CuLEA5*. Genomic DNA (10 mg) was digested with *Bam*HI (B), *XhoI* (XI), and *EcoRI* (E). Size markers (kbp) are indicated at left. For northern blot, bottom panel represents ethidium bromide-stained gel used for monitoring equal amounts of total RNA loading. L, leaf RNA; Lanes 1-7, fruit RNA at seven stages of fruit ripening. RT-PCR was performed using total RNA samples from young (Y) and old (O) leaves.

gymnosperm groups, as occurred with the P-segments (Fig. 1B).

Because plant Lea5 proteins, including those of cotton (Galau et al., 1993) and spruce (GenBank No. L47602 and L47628), exist in two more isoforms encoded by different genes, our results suggest that *Citrus* has two more isoforms of Lea5. In cotton *Lea5*, its two isoforms – *Lea5-A* and *Lea5-D* – exist as approximately 11.4-kDa proteins, with substitutions in six amino acids of their polypeptides (Galau et al., 1993). Our Southern blot analysis of the *Citrus* genomic DNA revealed two more bands, thereby indicating that the gene is present in two more related sequences, including *CuLEA5* (Fig. 2A).

Most studies of the Lea5 gene have been focused on seed

development and stress tolerance. For example, the sweet orange *C-Lea5* gene is expressed at high levels in *Citrus* leaves treated with sodium chloride, and in *Citrus* seedlings exposed to stresses such as drought, heat, and cold (Naot et al., 1995). However, to understand the physiological functions of Lea5, it is necessary to conduct more promoter and expression analyses. Here, when *CuLEA5* expression was investigated via northern blots, using total RNA prepared from leaves and fruits (Fig. 2B), its transcripts were detected in both tissue types. Furthermore, transcript levels were low in the immature stages, but increased with maturation. That is, mRNA levels were minimal at 53 and 67 DAF, but rapidly rose by 95 DAF, then increased through the developed stages (109, 123, 137, and 153 DAF). RT-PCR analysis for

-1480 -1395	ccccccccccccccccccacaaaaaaaaaaaaaaaaaa	ALFIN(-) <u>tgccactqcqaq</u> ctctt ggcatattcatcaatto	GAMYB(+) tt <u>qtta</u> tttgatga caacqqqaatataac	GAMYB(-) Mataa <u>taac</u> aaatatttagtt Mactgaattgaagetteegaa	ttaa atta
1210	WUS(-)			GAT	A (+)
-1310	gtgtttacacactcggatc <u>gcattaacqca</u> a	tggtaattggtaaagaa	gtactgettggttg	jatgatgacgatg <u>atqatqat</u>	taat
-1225 -1140	CRT-DRE(+) gg <u>oorgao</u> ttcatatgtyggcatccatgaacc <u>attaa</u> gagagacgtttagtgcaagaaagtag	ABRE (-) aagaga <u>cacqtaqaq</u> gt aaacttttttgaagtco	tacgacttacgatg xaagggaggccgccc	HAHB gttgccaattgttaage <u>cto</u> actgatcatggcccttgego	4 (+) <u>cata</u> ggcc
-1055	NUS(-) attatgaatcatcattgcgtcg <u>aqattaaca</u>	<u>cc</u> cgtcatcgcgatt <u>ca</u>	HVDRF1(+)	TEIL(-) aaaatatatat <u>atqttcat</u> o	caa <u>a</u>
-970	TEIL(-) GAMYB(+) <u>aqcttcat</u> gcaaacgt <u>qqacqqttqttattt</u>	c catacgetteg <u>gtetae</u>	RT-DRE(+) <u>xcqacqaqcc</u> acaaa	ctttagttccggcggcaaca	aatt
-885	ctcttacgactctctgtgtctctccata <u>tat</u>	GATA(+) <u>ttqataaaataqaq</u> caq	SBF-1(+,-) ttatatttaaatat	SP8BF(+) <u>ta</u> agaa <u>qatactattaaaa</u> t	taaa
-800	SP8BF(+), SBF-1(-) attaga <u>gttactattttaactacagt</u> aattt	CAAI ttgtg <u>gtcattgqtta</u>	!(-) Maagaatttattaat	WUS(+),SBB ttgcataacaat <u>ataattaa</u>	F-1(- <u>atqt</u>
-715	ATHB5(+) ctatataaaaattgtgc <u>tcaataattca</u> at <u>t</u>	SBF-1(-) <u>staatttaattatcaa</u> a	tactcataaatta <u>t</u>	GT-1(-) GT tqaqaataactctaaattaa	T-1(- acca
-630	SP8BF(+) <u>catc</u> cttttatatgatagct <u>tatactatatt</u>	<u>ca</u> atgtggtgtgtgcgt	SBF-1(+) t <u>ttcactattaaaa</u>	SBF-1(+) taattettaaaataqageea	aaat
-545	cagaatttattcaacatctaatcttttttt	ATHB5(+,-) aa <u>acaattattaa</u> tatt	tctcgaacttttct	tootttoaaaggaatattoa	aaca
-460	ATHE5(+) ccaactcaateett <u>eeaataatttt</u> aaaaaa	ataaaacgaaagtatca	aattaatcgcatga	GT-1(-),GT-3 atacttete <u>aaacaqttac</u>	3A(+) <u>cqtt</u>
275				GT-1(+)	
-375	attrattatttgcataaaagatgacaagag	jaaagaatttgaaatt	aggaaaagatat <u>tt</u>	conqueraaquergateet	6- 10 10 10
-290	GAMYB(-), NUS	(~) GATA(+) aaagagataaqtttcaa	aggaaaagatat <u>tt</u> HAHB4(+) g <u>atgattattaa</u> at	aacgacattagtggccc <u>tqa</u>	<u>aaaa</u>
-375 -290 -205	GAMYB(-), NUS tgtctatgaccoggacc <u>ocqattaacaqacc</u> GATA(-) SP8BF(+), <u>attatcttqqa</u> tcagatatc <u>aqtactataat</u>	(-) GATA(+) aaagagataaqtttcaa AAHB4 aataataacattgtgtc	aggaaaagatat <u>tt</u> HAHB4(+) g <u>atqattattaa</u> at aggattogatgaat	aacgacattagtggccc <u>tqa</u> CARE(-) GBE c <u>atcaactcctaaqa</u> agagc	<u>aaaa</u> F1(+) qtqt
-375 -290 -205 -120 -35 50 135 220	GAMYB(-), NUS gatadagaccoccattaacaagac GATA(-) SP8BF(+), attatcttqqatcagatacagacc GATA(-) SP8BF(+), attatcttqqatcagatacagatactataat ABF1(+) Cagattactcagacaagatctttgtcttct; ctcttTATAAataacagacaagatctttgtcttct; CTACTGCAGGAGCAGAATCTTTCTTCTTCTG TATAAGTAAATAAAACAATGGCCGCGCTCTTC M A R C L TAGCAGgtctgtttacagaactttttctct	(-) GATA(+) <u>maagagataaqtttcaa</u> . HAHB4 <u>mataataa</u> cattgtgtc <u>gaaagaacggcccaacc</u> .tctTcctcacaagcac MGTAGCACTGTACAGA TTCAAAGCTAGCACTG F K A K L .tctctctctctctttg	HANB4(+) gatgattattaaat aggattogatgatat aggattogatgaat agoacaaagagaGG AACATCCCAGCTAG TTCTFGCTCCTGTT L L A P V ggttttttgaaa	aacgacattagtggcoct <u>qa</u> CARE(-) GBE C <u>atcaactcctaaqaagago</u> GAGCOGCTAGCTTTTGCCPA CTATCTTCATATCTATTAAG GCTGATGGCATCTCGCTCTC A D G I S L S tgataattttatcttgtttc	aaaa F1(+) gttt ATAG GTTC CCAT S I cttg
-290 -205 -120 -35 50 135 220 305 390 475 560 645 730 815 985	GAMYB(-), NUS tgtotatgaccoggacc <u>ocqattaacaqacc</u> GATA(-) SP8BF(+), <u>attatcttqqa</u> tcagatatc <u>aqtactataat</u> ABF1(+) <u>caqatta</u> ctcagacaagatottttgtcttct ctcttTATAAataacagaaacctaaagcc CTACTGCAGGAGCAGAATCTTCTTCTTCTGC TATAAGTAAATAAACAATGGCCCGCTGTCT M A R C L TAGCAGgtctgtttacagaacatcatcata acattatgagagatcagtcattctcta s R ttctacggttgtgtogcttgatctcatcata acacaaaacaacgtagctttctgtaact gatcagattcatcgagggttagaacataaaa gctaaattattttgtaccttatcttct tagatcattagttcgttgtccagaactt tagatcattgttcgttgtcacatggaat tcccttaggatcattcgttcacatggaat	(-) GATA(+) <u>maggagataagtttcaa</u> <u>AABB4</u> <u>mataataa</u> cattgtgto <u>tetTectcacaagcac</u> <u>tetTectcacaagcac</u> <u>tetTectcacagcac</u> <u>tetCtcctctctttg</u> <u>tetctctctctttg</u> <u>tetctctctctttg</u> <u>tetcttcttttttt</u> <u>tegcatgtgatttga</u> <u>tetgaatcaggggg</u> <u>tetataataaggggc</u> <u>ttataagaac</u> <u>ttgagacgcataatt</u> <u>tgcaacttttgatg</u> <u>ttataataatttct</u>	HANB4(+) gatqattattaa aggattogatgat caastocaacaagt caastocaacaagt AACATOCCAGCTAG TTCTTGCTCCTGTT L L A P V ggtottotttgaa tatattgttott tocttostaactta tocgacttogtagt ccaasaggtaactt caatagaacettgg agtaatottttot aaatggacagGAGA	acgacattagtggcoctga CARE(-) GBE Catcaactoctaaqagagaco GAGCOGCTAGCTTTTGCCTA CATCTTCATATCTATTAAQ GCTGATGCCATCTCGCTCTC A D G I S L S tgataattttatottgtttc gatcaatctaacgattcaa acaacaacggattcaat cotcogcaatactttta agattgatattatagggttt egtcacttgtogtgcaagct ttgttatatattotcttt GGATATGCTGCACCAGCACC	aaaa F1(+) utgt ATAG STTC CCAT S II agtg aaga aagg ttag tagc tagc tagc tagc
-375 -290 -205 -120 -35 50 135 220 305 390 475 560 900 985 1070 1155 1240	GAMYB(-), NUS $tgtctatgacceggacceceqattaacaqacce GATA(-) SP8BF(+), attatcttqqatcagatacaqaccetacaqaccetacaqaccetacaqaccetacaqaccetacaqacaqatctttgtcttcts ctcttTATAAtaacaqaaacacctaagaccetacaqaccetacaqaccetacaqacaqaccetacaqacaagaccetacaqacaqatctttgtcttcs CTACTGCAGGAGCAGAATCTTTCTTCTTCTTCTG TATAAGTAAATAAAACAATGGCCGCGCTGTTCT M A R C L TAGCAGgtctgtttacaqaactetttttcts s R ttctacgttgtgtgcgttgatctatcatcata acactatagaagagatcaqtcattctgaaact gataggattcatcgggttagaacataaaa gctaaatagttatttgtttcacqgttgcagaactt tagatcattagtacacctggttgcagaactt tagatcatagatagttgttgcagaact gctaaatagttattgttctcgttacaga GGGGACCATATCTCGGACTGGTATCATGGAGG G T I S R T G I M E GCCCCAAAGTCACGGGCACATAGGGAGAGGGTGT$	(-) GATA(+) <u>maagagataaggtttgaatt</u> <u>maagagataagtttcaa</u> <u>AABB4</u> <u>maagaacggcccaacc</u> <u>itetTectcacaagcac</u> <u>itetTectcacagcac</u> <u>itetTectcacagcac</u> <u>itettcctctctttg</u> <u>itgaatcctagctgtt</u> <u>itgaatcctagctgtt</u> <u>itgaatcctagctgtt</u> <u>itgaatccttagctgtt</u> <u>itgaatcctttgaagggg</u> <u>itataatataagggac</u> <u>itataaacgagggg</u> <u>itataatataagggac</u> <u>itataaacgagggg</u> <u>itataatataagggac</u> <u>itataaacgagggg</u> <u>itataaacgatttgaatcctttaa</u> <u>itataataaattctt</u> <u>itataataattctt</u> <u>itataataatttct</u> <u>itataaacttttgaacgccc</u> <u>itataaacttttgaacgccc</u> <u>itataaacttttgaacgcc</u> <u>itataaacttttgaacgcc</u> <u>itataaaacttttgaacgcc</u> <u>itataaaacttttgaacgcc</u> <u>itataaaacttttgaacgcc</u> <u>itataaaacttttgaacgcc</u> <u>itataaaacttttgaacgcc</u> <u>itataaaacttttgaacgcc</u> <u>itataaaacttttgaacgcc</u> <u>itataaaacttttgaacgcc</u> <u>itataaaactttt</u> <u>itataaaactttt</u>	HANB4(+) gatgattattaaat caaggattogatgaat caaggattogatgaat caaggattogatgaat caaggattogatgaat caaggacaaagagaGG AACATCCCAGCTAG TTCTGGTCGTGTGGTGAGAGAGA ggttatttgtttott toottoataactta ttogatttogtatgaa coaaggaacttgg coaaggacacttgg coaaggacacttgg coaaggacacttgg coaaggacacttgg coaaggacacttgg GGGGTGAGAGAGAGG A V R E GAGATGATCAGCGGT	aacgacattagtggcoct <u>qa</u> CARE(-) GBE C <u>atcaactcctaaqaagaco</u> GAGCCGCTAGCTTTTGCCTA CTATCTTCATATCTATTAAG GCTGATGGCATCTCGCCTCT A D G I S L S tgataatttatcttgtttc gatcaatctaaacatttcta accaacaaacgaattcaaa tgaaaataaaat	aaaaa F1(+) gttt ATAG STTC CCAT CCAT agtg aaaaa aaat ttgg aaaaa aaat ttgg tagc ttgg tagc ttgg ttgg
-375 -290 -205 -120 -35 50 135 220 305 390 475 560 645 7300 985 1070 1155 1240 1325 1240 1325 1410 1495 15805	$\label{eq:GAMMB} (=), & & & & & & & & & & & & & & & & & & &$	<pre>(-) GATA(+) aaagaqataaqttttaa (-) GATA(+) aaagaqataaqttttaa , HAHB4 aataataacattgtgtc gaaagaacggcccaacc tctTcctcacaagcac rAGTAGCACTGTACAGA TTCAAGCTGTAAGCTTC F K A K L tctctctctctctttg ttgaatccttgqdgtt tagtttttcttttttt tgcaactatgagggc taataaaattcttt GGCaAGATCGGCGCGGCC C B N R A V CGCGATGGCAGATCCT FCTTGCTGTGTGTTT Tagtaagaattcctgt tgggaccatgtgtagca tagcacgtagtat tagcaacgtagtat</pre>	Aggaaaagatat <u>t</u> HAHB4(+) Igatqattattaaat aggattogatgaat aqoacaaagagaGG AACATCCCAGCTAG TTCTTGCTCCTGTT L L A P V Iggtcttcttgaaa tatattgtttctt tccttcataactta tcgacttcgtagg ccaaaaggaaggtaact tctgacattgttctt tcctfcataactta aatggacagGAGA B CGGGTGAGAGAGCAG E I D P A TTGCTGTTAGCGGT GATATCAACTGTA GATATCAACTGTA Attgcaaaattact gataataataataa taacattgcaaagtaca gataattgcaaagtac	aacgacattagtggcoct <u>qa</u> CARE(-) GBE C <u>atcaactoctaaqaagag</u> CGCCtotgcacagtttotog GAGCOGCTAGCTTTTGCCTA CTATCTTCATATCTATTAAG GCTGATGGCATCTCGCTCTC A D G I S L S tgataattttatottgttto gatcaatctaaacattcta acaacaaacgaattcaaa agatggacggtotcaat cotcogcaaatactttgga agatggtatattaggat cgtcacttgtogtgcaggt ttgttatataattttta GGTATGGCACCTCTCTGG D S G A S S A AGAGCTTCGAGAGATCTGT . E L R E I L GCTTTTAACTTTTAATGGGG TTGACAACTTTTAATGGGG tttaataggggtatgctg taataaatcaacggcoto ttgtagaatcagtcaacg	aaaa gtitt ATAG GTTT S I CCCAT S I CCCAT S I CCCAT S I CCCAT CCCAT CCCAT CTGG A AGGT CATT GATT CATT CATT CATT CATT

Figure 3. Nucleotide sequence and deduced amino sequence for *CuLEA5* genomic clone of *Citrus*. Uppercase letters represent exons; lowercase letters indicate introns, 5'-upstream, and 3'-downstream regions. Putative TATA box and transcription start site "T" is represented by bold, uppercase letters between –35 and +1 positions. Underlined letters in bold indicate P-segment. Predicted promoter elements relating to hormonal and stress responses are represented both by underlines and letters above: ALFIN1, zinc-finger protein binding site, regulates salt tolerance; GAMYB, GA-regulated myb binding site; WUS, homeodomain protein WUSCHEL binding site; GATA, class I GATA factor binding site; CRT-DRE, C-repeat/dehydration response element; ABRE, ABA response element; HAHB4, sunflower homeodomain leucine-zipper protein Hahb-4 binding site; HVDRF1, barley dehydration-response factor 1 binding site; TEIL, tobacco EIN3-like protein binding site; SBF-1, SBF-1 transcription factor binding site; GT1, GT1-box; CARE, GA-inducible CAACTC regulatory element; ABF1, ABA inducible transcriptional activator binding site; GBF1, bZIP G-box.

leaf development showed similar expression patterns, with transcripts being two to three times higher in old leaves than in young leaves (Fig. 2C).

Probes make it possible to detect expression for several homologous genes. As determined from our Southern blot analysis, those patterns suggested that CuLEA5 has important roles in the development or ripening of *Citrus* fruit, such as conferring protection again internal and external stresses. It is also likely that *CuLEA5* expression patterns are influenced by environmental conditions. For example, the immature stages (53 and 67 DAF) of fruit development coincided with the heavy rainy season during Korean summers,

while other stages corresponded to the dry period leading into autumn, when temperatures also were decreasing. Therefore, these variations in water conditions and weather may have affected expression. Naot et al. (1995) have reported that the expression of *C-Lea5* in orange is enhanced by drought and cold. Thus, all of these results demonstrate that environmental conditions and particular developmental stages are key factors that regulate *Lea5* expression in *Citrus*.

In the case of pineapple, the *Lea5* gene, *PAE122*, is expressed only in the aerial parts – including the green leaves, shoots, and stems – but not in the fruits or at any stage of ripening (Neuteboom et al., 2002). This suggests that expression patterns vary with plant species.

Gene expression is regulated mainly by the promoter region. To characterize this region in *CuLEA5*, we first isolated a 3.2-kbp genomic DNA (GenBank accession number DQ424891) containing the *CuLEA5* gene from the *Citrus* genomic library (Fig. 3). Comparison with the gene sequence of the cDNA revealed its exon/intron organization. The isolated *CuLEA5* gene comprised two exons interrupted by one intron (Fig. 3), a structure similar to that of cotton *Lea5* (Galau et al., 1993). Furthermore, the transcription start site and putative transcription binding sites in the *CuLEA5* promoter were inferred from the Neural Network Promoter Prediction (http://www.fruitfly.org/seq_tools/promoter.html) and the Genomatrix Matlspector (http://portal1.0.genomatix.de/index.html).

Several potential regulatory elements were found between positions -1480 and 1, and a putative TATA box (TATAAA) was identified at position -30. The promoter region contained various putative stress- and hormonal-responsive elements (Fig. 3).

The promoter region of *CuLEA5* shared no significant similarity with that of the cotton *Lea5* genes, with some exceptions. For example, the prediction of *cis*-elements using the cotton *Lea5* promoter (GenBank accession No. X54448) indicated that the latter had putative TATA and CAAT boxes like those of the *CuLEA5* gene promoter. Moreover, the putative GAMYB, TEIL, GT-1, GATA, SBF-1, SP8BF, WUS, ATHB1, ATHB5, and HAHB4 elements also were found in the cotton *Lea5* promoter region, albeit at different positions (data not shown). The existence of these shared features suggests a common regulatory mechanism for gene expression that is mainly influenced by plant hormones and stresses (Kim et al., 2003).

In addition, the different expression patterns for *CuLEA5* genes, which depend upon tissue type, developmental stage, or environmental stress, may be controlled by various regulatory elements in the promoter region. Therefore, to investigate the hormonal- and stress-responsive regulation of this expression, one must examine the structure of the *CuLEA5* gene promoter. Such an analysis, using transgenic plants, is in progress.

ACKNOWLEDGEMENTS

This work was supported by a grant (20050301034400) from the BioGreen 21 Program of the Rural Development

Administration, Korea, and by the Technology Development Program for Agriculture and Forestry, Ministry of Agriculture and Forestry, Korea.

Received May 15, 2006; accepted August 24, 2006.

LITERATURE CITED

- Artus NN, Uemura M, Steponkus PL, Gilmour SJ, Lin C, Thomashow MF (1996) Constitutive expression of the cold-regulated Arabidopsis thaliana COR15a gene affects both chloroplast and protoplast freezing tolerance. Proc Natl Acad Sci USA 93: 13404-13409
- Baldwin EA (1993) Citrus fruit, *In* GB Seymour, JE Taylor, GA Tucker, eds, Biochemistry of Fruit Ripening. Chapman and Hall, New York, pp 107-150
- Bernacchia G, Furini A (2004) Biochemical and molecular responses to water stress in resurrection plants. Physiol Plant 121: 175-181
- Burns WC, Maitra N, Cushman JC (1996) Isolation and characterization of a cDNA encoding a LEA5-like protein from soybean (U66316)(1)(PGR96-03). Plant Physiol 112: 1398
- Close TJ (1997) Dehydrins: A commonality in the response of plants to dehydration and low temperature. Physiol Plant 100: 291-296
- Dellaporta SL, Wood J, Hicks JBA (1983) Plant DNA minipreparation: Version II. Plant Mol Biol Rep 1: 19
- Galau GA, Wang HYC, Hughes DW (1993) Cotton *Lea5* and *Lea14* encode atypical late embryogenesis-abundant proteins. Plant Physiol 101: 695-696
- Giovannoni JJ (2004) Genetic regulation of fruit development and ripening. Plant Cell 16: S170-S180
- Ha YI, Lim JM, Ko SM, Liu JR, Choi DW (2006) Sequence variability and expression characteristics of the ginseng (*Panax ginseng* C.A. Meyer) dehydrin gene family. J Plant Biol **49**: 205-211
- Hara M, Terashima S, Fukaya T, Kuboi T (2003) Enhancement of cold tolerance and inhibition of lipid peroxidation by citrus dehydrin in transgenic tobacco. Planta 217: 290-298
- Hara M, Wakasugi Y, Ikoma Y, Yano M, Ogawa K, Kuboi T (1999) cDNA sequence and expression of a cold-responsive gene in *Citrus unshiu*. Biosci Biotechnol Biochem 63: 433-437
- Hwang EW, Kim KA, Park SC, Jeong MJ, Byun MO, Kwon HB (2005) Expression profiles of hot pepper (*Capsicum annuum*) genes under cold stress conditions. J Biosci 30: 657-667
- Jackson S, Gascon J, Carrera E, Monte E, Prat S (1997) Cloning and expression analysis of a gene that shows developmental regulation upon tuberization in potato. Plant Mol Biol 33: 169-174
- Kim IJ, Ko KC, Kim CS, Chung WI (2001a) Isolation and expression patterns of a cDNA encoding phytoene synthase in *Citrus*. J Plant Physiol 158: 795-800
- Kim IJ, Noh SJ, Lee BH, Jo J, Kim YS, Chung WI (2001b) Molecular characterization of cDNA clones for ADP-glucose pyrophosphorylase from *Citrus*. Biochim Biophys Acta 158: 324-328
- Kim SH, Lee JR, Kim SR (2003) Molecular characterization of a fruit-preferential thaumatin-like gene from apple (Malus domestica cv. Fuji). J Plant Biol 46: 52-58
- Naot D, Ben-Hayyim G, Eshdat Y, Holland D (1995) Drought, heat and salt stress induce the expression of a citrus homologue of an atypical late-embryogenesis *Lea5* gene. Plant Mol Biol 27: 619-622
- Ndong C, Danyluk J, Wilson KE, Pocock T, Huner NPA, Sarhan F (2002) Cold-regulated cereal chloroplast late embryogenesis abundant-like proteins: Molecular characterization and functional analyses. Plant Physiol 129: 1368-1381

- Neuteboom LW, Kunimitsu WY, Webb D, Christopher DA (2002) Characterization and tissue-regulated expression of genes involved in pineapple (*Ananas comosus* L.) root development. Plant Sci 163: 1021-1035
- Park BJ, Liu Z, Kanno A, Kameya T (2005) Transformation of radish (*Raphanus sativus* L.) via sonication and vacuum infiltration of germinated seeds with *Agrobacterium* harboring a group 3 *LEA* gene from *B. napus*. Plant Cell Rep 24: 494-500
- Porat R, Pasentsis K, Rozentzvieg D, Gerasopoulos D, Falara V, Samach A, Lurie S, Kanellis AK (2004) Isolation of a dehydrin cDNA from orange and grapefruit citrus fruit that is specifically induced by the combination of heat followed by chilling temperatures. Physiol Plant 120: 256-264
- Porat R, Pavoncello D, Lurie S, McCollum TG (2002) Identification of a grapefruit cDNA belonging to a unique class of citrus dehydrins and characterization of its expression patterns under temperature stress conditions. Physiol Plant 115: 598-603
- Shao HB, Liang ZS, Shao MA (2005) LEA protein: Structure, functions and gene expression. Colloids Surf B Biointerfaces 45: 131-135
- Sivamani E, Bahieldinl A, Wraith JM, Al-Niemi T, Dyer WE, Ho TD, Qu R (2000) Improved biomass productivity and water use efficiency under water deficit conditions in transgenic wheat constitutively expressing the barley *HVA1* gene. Plant Sci 155:

1-9

- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The ClustalX windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucl Acids Res 25: 4876-4882
- Verwoerd TC, Dekker BM, Hoekema A (1989) A small-scale procedure for the rapid isolation of plant RNAs. Nucl Acids Res 17: 2362
- Wang W, Vinocur B, Altman A (2003) Plant responses to drought, salinity and extreme temperatures: Towards genetic engineering for stress tolerance. Planta 218: 1-14
- Wise MJ (2003) LEAping to conclusions: A computational reanalysis of late embryogenesis abundant proteins and their possible roles. BMC Bioinformatics 4: 52
- Xu D, Duan X, Wang B, Hong B, Ho HD, Wu R (1996) Expression of a late embryogenesis abundant protein gene, *HVA1*, from barley confers tolerance to water deficit and salt stress in transgenic rice. Plant Physiol 110: 249-257
- Yamamoto KT (1994) Further characterization of auxin-regulated mRNAs in hypocotyl sections of mung bean [*Vigna radiata* (L.) Wilxzek]: Sequence homology to genes for fatty-acid desaturases and atypical late-embryogenesis-abundant protein, and the mode of expression of the mRNAs. Planta **192**: 359-364