Plant Molecular Biology **26**: 617–630, 1994.

Structure and promoter analysis of an ABA- and stress-regulated barley gene, *HVA1*

Peter F. Straub [1], Qingxi Shen [1] and Tuan-hua David Ho*
Plant Biology Program, Department of Biology, Division of Biology and Biomedical Sciences, Washington University, St. Louis, MO 63130, USA (author for correspondence)*

Received 20 April 1994; accepted in revised form 6 July 1994

Key words: ABA, barley, gene expression, *Hordeum vulgare*, phylogeny, stress

Abstract

A single-copy barley gene, *HVA1*, encoding a class 3 late embryogenesis-abundant protein, can be induced by either treatment with abscisic acid (ABA) or by stress conditions such as drought, cold, heat and salinity. We have isolated an HVA1 genomic clone containing about 400 bp of 5'-upstream sequence, a single 109 bp intron, and the full coding sequence. Linker scan mutagenesis and transient expression studies were used to test the function of four *HVA1* promoter elements conserved in ABA-responsive genes. Mutations in two of these elements, the C box and the putative ABRE 1 (ABA-responsive element) containing an ACGT core, resulted in no significant change in transcription level or ABA induction. In contrast, mutations of the other two elements, putative ABRE 2 & 3 cause the level of transcription to drop to 10–20% of that obtained with the wild-type promoter indicating that the high level of expression of *HVA1* is dependent on both pABRE 2 & 3. Interestingly, despite their low level of expression, the mutated promoters still gave more than 20-fold induction in response to ABA treatment. We suggest that the ABA induction of barley *HVA1 gene* is governed by a complex consisting of pABRE 2 & 3 working together to regulate the absolute level of expression, and either of these elements or a possible third element may regulate ABA inducibility. Phylogenetic analysis by parsimony indicates that the barley *HVA1* and wheat *pMA2005* sequences share a recent common ancester. These two genes are closely related to the carrot *Dc3* and cotton *D-7* genes with which they share a similar structural gene organization.

Introduction

The phytohormone abscisic acid (ABA) mediates seed development and dormancy, and adjustment to environmental stresses. In order to study the molecular mechanism of ABA action, ABA-responsive genes have been isolated from a variety of plants, including barley, wheat, rice, maize, cotton, rapeseed and soybean (see [43] for review). Many of these genes are termed *Lea* (late

[1] P. Straub and Q. Shen have made equal contributions to this work and should be considered as co-authors.
The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession number 78205.

embryogenesis-abundant) genes which express at late stages of seed development, and the other are designated as *Rab* (responsive to ABA) genes which express in stressed vegetative tissues. Other ABA-responsive genes include those encoding a barley aldose reductase [3], a maize RNA-binding protein [34], a putative transcription factor [47] and a putative protein kinase [1]. The *Lea* genes have been divided into three groups based on common amino acid sequence domains, and members in each group are probably similar in function [14]. The barley *Lea* gene we have studied, *HVA1* [28], belongs to the group 3 which encompasses other members such as cotton *D-7*, and *D-29* [2, 15], *Brassica pLEA76* [22], carrot *Dc8* and *Dc3* [19, 40], soybean *pmGM2* [29], and wheat *pMA2005* and *pMA1949* [9, 10]. These proteins are characterized by 11 amino acid tandem repeats which could form a probable amphiphilic α-helical structure that presents a hydrophilic surface with a hydrophobic stripe [13]. These proteins have been hypothesized to reorient cellular water molecules, and bind ions to ameliorate the consequences of water removal during desiccation. The induction of LEA proteins is highly correlated with the maturation and desiccation of the embryo as it completes the final steps in a developmental program designed to maintain dormancy and protect the embryo during seed dispersal.

In parallel to the studies of the physiological roles of ABA, efforts are being made to investigate the molecular mechanism of ABA action, including the definition of ABA-responsive elements (ABRE) and the *trans*-acting factors that interact with ABREs. It is reported that a 75 bp fragment of the ABA-inducible wheat *Em* gene, when fused in either direction to a truncated (−90) 35S promoter, gave a greater than 10-fold induction of GUS activity in the presence of ABA [21]. A leucine-zipper protein, EmBP-1, is shown to bind the ABRE sequence (CACGTGGC) in this 75 bp region. Mutation of this sequence abolishes the binding of EmBP-1 to this sequence as well as the ABA responsiveness of the chimeric promoter [21]. Skriver *et al.* [44] demonstrated that six copies of the sequence GTACGTGGCGC conferred ABA inducibility to a truncated 35S minimal promoter. In a similar study with barley *HVA22*, we have shown that at least three elements, two located in the promoter and one in the first intron are essential for the high-level expression of this gene in response to ABA treatment [42]. Recently, Michel *et al.* [36] report that two promoter regions of an ABA- and stress-regulated *Craterostigma plantagineum* gene appear to be involved in the ABA response.

The expression of *HVA1* is tissue-specific and temporally regulated in developing seeds, with its mRNA detected in aleurone layers and embryos, but not in the endosperm [27]. In vegetative tissues, *HVA1* is also developmentally and organ-specifically regulated by ABA and drought treatment. Accumulation of its mRNA is also induced by salt, cold and heat stress [27]. The purpose of this study was to isolate and characterize the HVA1 genomic sequence and investigate the role of upstream sequence elements in ABA induction of the *HVA1* gene. We have found that two putative ABRE sequences are necessary for the high-level expression of *HVA1*, yet mutations in either of these sequences do not seem to affect ABA inducibility. An additional objective of this work was to place the *HVA1* gene sequence within a phylogenetic tree of related *Lea 3* gene sequences.

Materials and methods

Chemicals

Unlabeled deoxynucleotide triphosphates and random hexamers were acquired from Boehringer Mannheim Biochemicals, Indianapolis, IN. Deoxycytidine α-[^{32}P]-triphosphate (3000 Ci/mmol), guanidine α-[^{32}P] triphosphate (3000 Ci/mmol) and [^{35}S]-deoxyadenosine 5′(α-thio)-triphosphate (1185 Ci/mmol) were purchased from NEN Research Products, Boston, MA. Sequenase 2.0 sequencing kit, exonuclease III, and acrylamide were obtained from USB, Cleveland, OH. λZAPII, packaging extracts, BlueScript plasmid vectors and sequencing oligonucleotides

were obtained from Stratagene, La Jolla, CA. Restriction enzymes, T4 DNA ligase, polynucleotide kinase and Klenow were obtained from Promega, Madison, WI; New England Biolabs, Beverly, MA; and International Biotechnologies, New Haven, CT. Agarose, low-melting-temperature agarose, Tris, bis-acrylamide, urea and guanidine-HCl, were obtained from Fisherbiotech, Fairlawn, NJ. Unless specified all other reagents were acquired through Sigma Chemical Co., St. Louis, MO.

Treatment of plants

Barley seeds (*Hordeum vulgare* L. cv. Himalaya), obtained from the 1985 harvest at Department of Agronomy and Soils, Washington State University, Pullman, WA, were used throughout this study. Three-day-old seedlings were treated with the following conditions. For ABA treatment, the seedlings were sprayed and watered with 0.1 mM ABA for 24 h. For the drought treatment, the seedlings were dehydrated to 85% of original fresh weight and kept dehydrated for 24 h. For cold treatment, the seedlings placed on ice in a cold room (1 °C) for 96 h. For cold stress recovery, the cold-stressed plants were transferred back to room temperature (22 °C) for 14 h. For heat treatment, the seedlings were placed in a flask of 100% relative humidity and then transferred to a 37 °C water bath for 2 h. For NaCl treatment, seedlings were watered with 0.17 M NaCl for 72 h.

Isolation of total RNA and northern blot analysis

Total RNA was isolated using the guanidine HCl method as described previously [4]. Ten μg of total RNA from each sample was electrophoresed in formaldehyde-agarose gel and blotted on a GeneScreen membrane (New England Nuclear, Boston, MA) following Maniatis *et al.* [35] and the manufacturer's instructions. Hybridization and washing were performed according to the method of Church and Gilbert [7]. The probe used was random-primed [^{32}P]-labeled cDNA of HVA1 [28]. For autoradiography, the membrane was exposed to a Kodak XAR film at −80 °C.

Isolation of genomic DNA

Shoots of two-week-old barley seedlings were frozen in liquid nitrogen, ground with dry ice in a coffee grinder and genomic DNA was isolated by a procedure modified from Sutton [45]. Briefly, 50 g of ground plant material was homogenized in 50 ml of extraction buffer (0.2 M Tris-HCl pH 8.5, 10 mM EDTA, 1% SDS). The homogenate was then extracted with one volume (100 ml) of cold phenol/chloroform/isoamyl alcohol (25:24:1) by shaking on ice (30 min) and centrifuged at 8000 × *g* in a Sorvall GSA rotor. The aqueous layer was removed, and the phenol layer back-extracted with 0.2 vol of buffer. The combined aqueous extract was adjusted to 0.5 M $NaClO_4$ and again extracted with an equal volume of cold phenol/chloroform/isoamyl alcohol (25:24:1) by shaking on ice followed by centrifugation at 8000 × *g*. The aqueous phase was then precipitated with 2 vol of ethanol and the DNA was recovered by spooling on a glass rod. The spooled DNA resuspended in 20 ml TE buffer dialyzed against TE. The dialyzed DNA was then treated with RNase A (boiled) at 50 μg per ml for 60 min at 37 °C. The DNA was then made to 0.15 M with NaCl and extracted twice with an equal volume of chloroform/isoamyl alcohol (24:1) on ice for 30 min. The aqueous layer was recovered after centrifugation at 8000 × *g* and the organic layer back extracted with 0.2 vol of TE containing 0.15 M NaCl. The combined aqueous extract was precipitated with 2 vol of ethanol and the DNA recovered by spooling on a glass rod. Spooled DNA was resuspended in 10 ml of TE and extensively dialyzed against TE.

Southern blot analysis

Ten μg of genomic DNA was digested with a number of restriction enzymes and electrophore-

sed through 0.8% agarose gel. After the acid depurination and neutralization, the gel was blotted to nylon filters (GeneScreen; NEN, Boston, MA) with 10 × SSC. Filters were UV crosslinked for 2 min at 254 nm, prehybridized, hybridized and washed at 65 °C by the method of Church and Gilbert [7]. Probes were prepared with random hexamers labeling [17] from the entire *Eco* RI cDNA insert, an *Ava* II/*Eco* RV 3′ fragment or an *Ava* II/*Bam* HI 5′ fragment of HVA1 cDNA [28].

Isolation of HVA1 genomic clone

A partial genomic DNA library was prepared as following. *Bam* HI-digested genomic DNA was size selected by electrophoresis in 0.6% agarose in TAE buffer and the 2.5–4.0 kb fraction was electroeluted and purified. Sized DNA fragments were filled in with dATP and dGTP using the Klenow enzyme. λZAPII (Stratagene, La Jolla, CA) DNA was digested with *Xho* I and filled in with dCTP and dTTP to create cohesive termini with the insert DNA while preventing self-ligation. Insert and vector DNA were ligated with T4 DNA ligase at a 1:1 molar ratio, 0.5 μg phage arms to 0.05 μg insert, and packaged according to the manufacturer's instructions.

The library was screened by plating with *E. coli* strain ER1647 at a titer of 6×10^3 plaques per 150 mm plate. The *Ava* II/*Eco* RV 3′ fragment of the HVA1 cDNA [28], was used as a probe. A total of 1.3×10^5 plaques were screened in duplicate resulting in two positive plaques. After the secondary screening, positive clones were recovered in the BlueScript SK(−) plasmid by *in vivo* excision method as instructed by the manufacturer (Stratagene, La Jolla, CA).

To sequence the genomic clone, Exonuclease III and S1 nuclease were used to create a unidirectional nested set of overlapping deletions on both strands [23]. The DNA sequence was determined by the Sanger dideoxy chain termination method [39] using the Sequenase 2.0 kit (United States Biochemical, Cleveland, OH).

Gene copy number reconstruction

Single- and multiple-copy reconstructions were calculated based on the size of the haploid genome being 5.5×10^{-12} g. To purify the cloned insert DNA it was necessary to cut with *Bst* XI which resulted in the loss of 145 bp of 5′ sequence since the cloning procedure destroyed the original flanking *Bam* HI restriction sites. This loss was included in all subsequent calculations. For this cloned insert, 11.1 pg was calculated to be the single copy equivalent of HVA1 DNA present in 10 μg of genomic DNA. Electrophoresis, blotting, and hybridization were as above in the Southern analysis.

Primer extension

A 30 base oligonucleotide (5′-TGGTA GCTCC CCTGG TTCTG GTTGG AGGCC-3′, complementary to the sequence between + 134 and + 95, Fig. 4) was synthesized at the DNA Chemistry Facility, Dept. of Biology, Washington University. Oligonucleotide labeling with polynucleotide kinase, hybridization and reverse transcription were done essentially as described [20]. Total RNA (30 μg), isolated with guanidine-HCl from mature embryos and aleurone layers [4] treated with or without 0.01 mM ABA for 24 h, was hybridized overnight at 30 °C with the ^{32}P-labeled oligonucleotide. *E. coli* tRNA was also included as a negative control. Reverse transcription was conducted at 42 °C for 2 h. The primer extension oligonucleotide was used as a sequencing primer to generate a size standard for comparison with the extension product. Extension products and size standard were electrophoresed and visualized by autoradiography as for DNA sequencing.

Phylogenetic analysis of Lea 3 coding sequences

DNA sequence was obtained for the Lea 3 genes presently available from the databases, EMBL and GenBank, and from published sources: cotton *D-7*, and D-29 [2, 15], *Brassica pLEA76* [22], carrot *Dc8* and *Dc3* [19, 40], soybean GmPM2

[29] and wheat *pMA2005* and *pMA1949* [9, 10]. Since genomic sequences were not available for all the clones, we limited the analysis to DNA of the coding regions. The GCG routine 'Pileup' [11] was used for initial comparisons. To obtain a reasonable block of sequence for comparison, and since the clones are of varying size, the analysis was limited to the region of overlap present in all the clones with appropriate gaps for insertions/deletions. The Phylogenetic Analysis Using Parsimony (PAUP) version 3.0 [46] was used with a heuristic search strategy for this analysis. Statistical testing was conducted using a resampling bootstrap technique [16] for evaluating the accuracy of clades [18] and a g_1 statistic was used to evaluate the statistical confidence levels of the possible tree lengths based on skewness of a distribution of all possible tree topologies [25].

DNA constructions and transient expression

A *Sal* I site was introduced at +98 (Fig. 3) by PCR with the *Bam* HI-cut DNA of the HVA1 genomic clone as the template. The two primers used were the M13-20 (5′-GTAAAACGACGGCCAG-3′), which is a vector sequence, and the oligonucleotide (5′-AGGTC GACCTTCGTCTCACGATCTAGCT-3′), which includes the HVA1 sequence between +92 and +83 plus the *Sal* I site (underlined) and two extra nucleotides. The *Kpn* I/*Sal* I-digested PCR product was cloned into the corresponding sites of plasmid Bluescript SK⁻. Sequencing was performed to confirm the sequence of the resulting clone, designated PA1. The *Kpn* I/*Sal* I of PA1, containing the HVA1 promoter plus the 92 bp untranslated sequence of HVA1, was linked to the construct IGU, which contains the coding sequence for the *E. coli GUS* gene [31] with a modified ATG initiation codon [33], the intron 1-exon 2-intron 2 fragment and the 3′ region of HVA22 [42] to generate the construct WT. As a control, the construct IGU linked to a minimal Amy64 promoter was previously shown to not be inducible by ABA treatment in a transient assay system [41]. To make linker-scan mutants, the PA1 DNA was introduced into *E. coli dut⁻ ung⁻* strain (CJ236) and single-strand DNA was prepared. Three oligonucleotides, 5′-AAGGG GTGGC GCGTG gaatt ccgat CCGGC GAGCA CAAGT-3′ for pABRE1, 5′-GGTGC TCGTA CATGG gtgaa ttCac GGGAG GACAC GCGTT-3′ for pABRE2 and 5′-CGGGA GAGTC GCCGG tgaAt tcgta GCGGC GGTGC TCGTA-3′ for pABRE3, were synthesized with each containing a 9 or 10 bp mutation (lower case) and a *Eco* RI site (underlined) to facilitate screening. The oligonucleotides were hybridized to the single-strand DNA templates prepared from the construct PA1. *In vitro* DNA synthesis, ligation and transformation were conducted as described [32]. The mutated promoters were linked to the construct IGU to give the constructs LS-pA1, LS-pA2 and LS-pA3. The same procedure was performed with the oligonucleotide 5′-ACACG TTGCG CGGCG acgaa ttCtt ATGGC CGCCA CGTAG-3′ to produce the linker scan mutant at the C-box region, designated LS-Box C. The HVA1 promoter sequence in these mutants were sequenced to ensure no unwanted mutation was introduced. The junctions of all constructs were also confirmed by sequencing.

Embryoless half seed preparation, DNA construct bombardment, homogenization of the bombarded seeds, β-glucuronidase and luciferase assays were conducted essentially as described previously [42]. Briefly, test constructs, along with an internal control ubiquitin promotor/luciferase reporter construct [6], were coated on tungsten particles in a 1:1 molar ratio and introduced into aluerone tissue by microprojectile bombardment. Bombarded tissue was incubated with or without 0.02 mM ABA for 24 h, ground in buffer, and then aliquots assayed for GUS and luciferase activity.

Results

Regulation of HVA1 *gene expression by ABA and stress conditions*

Although the cDNA of HVA1 was first isolated from ABA-treated barley aleurone layers and classified as a LEA protein based on sequence

similarities, the expression of HVA1 can also be induced by various stress conditions in the vegetative tissues of three-day-old barley seedlings. Similar to the previous report by Hong *et al.* [27], the steady state level of HVA1 transcripts increased dramatically after treatment with ABA, drought, and cold (Fig. 1). The effect of salinity (watering with 0.17 M NaCl) and heat was much more modest compared to the other conditions. It is intriguing that heat induced the expression of HVA1 only in roots but not in shoots. The effect of cold induction of HVA1 could be readily reverted, no HVA1 transcripts were detected after the cold-stressed plants were transferred back to room temperature for 14 h. To investigate the mechanisms regulating the expression of HVA1, we have cloned and analyzed the promoter of this gene.

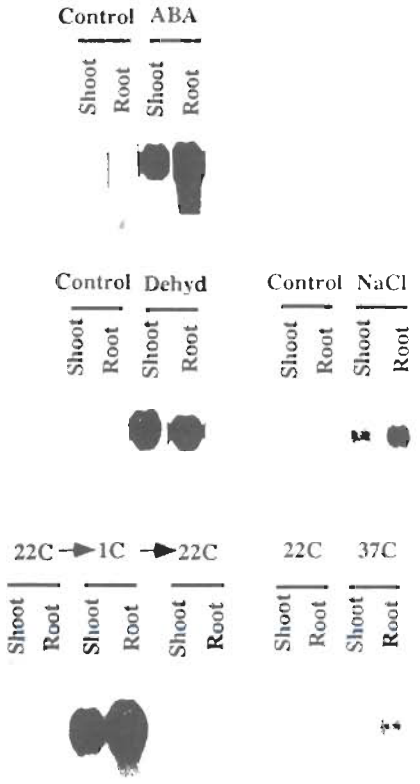

Fig. 1. Northern blot analysis showing ABA and stress induction of the *HVA1* gene in three-day-old barley seedlings. The plants were treated with ABA, drought (Dehyd), NaCl, cold (1C and recovery to 22C) and heat (37C). Detailed information of the treatments is described under the Materials and methods.

Structure, copy number, and transcription start site of the HVA1 *gene*

A 2.8 kb *Bam* HI fragment was identified, by genomic Southern analysis, that appeared to contain the entire gene plus a significant amount of 5′-upstream sequence (Fig. 2). A partial genomic library was constructed using *Bam* HI-digested DNA enriched for this fragment. Screening the partial genomic library resulted in two positive clones with different restriction maps. Initial DNA sequencing indicated that both of them matched the HVA1 cDNA but were cloned in opposite orientations.

Copy reconstruction of the *HVA1* gene from the cloned DNA indicated that this gene is probably present in a single copy in the barley haploid genome (Fig. 3). A single band in lanes loaded with the *Bam* HI-digested barley genomic DNA (lanes 1 and 2) hybridized to the HVA1 cDNA probe with the intensity similar to the single copy equivalent of the genomic clone (lane 3). The five-copy and ten-copy equivalent samples were included for comparison (lane 4 and 5). Not shown is the high molecular weight region where we noted some faint hybridization that could indi-

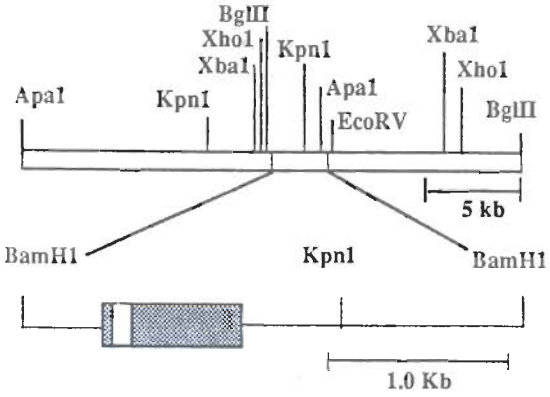

Fig. 2. Restriction map of HVA1 genomic clone. The boxed region is the HVA1-coding sequence with the hatched area being the coding region interrupted by the single intron.

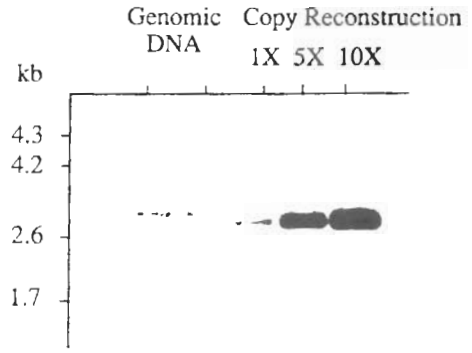

Fig. 3. Copy reconstruction experiment indicating that *HVA1* is a single-copy gene. The two left lanes are loaded with 10 μg of *Bam* HI-digested barley genomic DNA. The three right lanes contain aliquots of the HVA1 genomic clone insert calculated to approximate one (1 ×), five (5 ×), and ten (10 ×) copies of the *HVA1* gene sequence in 10 μg of total genomic DNA. The HVA1 insert used in the reconstruction is 165 bp smaller than the full length genomic clone because a convenient restriction site was unavailable to liberate and purify the entire fragment. The size difference results in the mobility difference on the Southern blot.

cate that there are distantly related sequences to HVA1 in the barley genome.

The HVA1 genomic clone was completely sequenced on both strands (Fig. 4) and matched the cDNA sequence perfectly except beyond nucleotide −72 of the cDNA published sequence [28]. This mismatch probably indicates that the cDNA had been subject to some rearrangement due to cloning. The numbering of the sequence is relative to the mapped transcription start site (+1, see below). The coding region is divided into two exons by a single 109 bp intron that is bordered by consensus intron splicing sequences [5]. The first exon is 20 amino acids in length and the second 193 amino acids. In the 5′ upstream region three putative ABA-responsive elements (pABRE) with ACGT cores were noted at −66, −98, and −244 as well as a sequence at −80 region, designated as C box, that is conserved in a number of ABA responsive genes. In the 3′-untranslated region a sequence similar to the consensus polyadenylation signal (AATAAA) was noted at +1070.

The transcription start site of HVA1 was mapped with the primer extension technique (Fig. 5). The site is designated as +1, which is 30 bp downstream from the TATA box and 12 bp from the CAAT box. No signal was detected in the non-ABA-treated sample (lane 6) or the *E. coli* RNA negative control (lane 7).

Phylogenetic analysis of the Lea 3 *gene family*

Comparison of nucleic acid and amino acid sequence of *Lea 3* genes has been extremely important in defining the conserved regions believed to be important to the function of these genes [14]. However, these comparisons have not attempted to use sequence variation to systematically relate the members of this gene family on a phylogenetic basis. For this purpose we used parsimony analysis to reconstruct the phylogeny and infer sequence evolution among the DNA of the various members of the *Lea 3* gene family. From the sequences input into the analysis, a consensus, majority rule, phylogenetic tree was drawn to represent the most parsimonious changes necessary to relate these sequences (Fig. 6). Numbers on the nodes of the tree indicate bootstrap values with 1000 replicates from repeated resampling of the population of all possible parsimonious tree configurations or topologies [18]. Bootstrap values ≥70 correspond to a probability of ≥95% that the clade supported by that node on the tree is real [24]. Except for two nodes, the bootstrap supports the phylogenetic inference of the parsimony analysis. The distribution of possible tree topologies is significantly skewed leftward, $g1 = -1.0509$ ($p < 0.01$), and indicates that the tree represents a true phylogeny rather than a random association of DNA sequences [25]. The monocot *HVA1* sequence and the monocot *pMA2005* genes formed a clade supported by a significant bootstrap value of 100 and were most closely related to the dicot carrot *Dc3* and dicot cotton *D-7* genes which also formed a clade supported by a significant bootstrap value of 100. These two clades were separated from the rest of the genes analyzed by a node in the tree with a significant bootstrap value of 85. The most distant groups from barley *HVA1* were the dicot

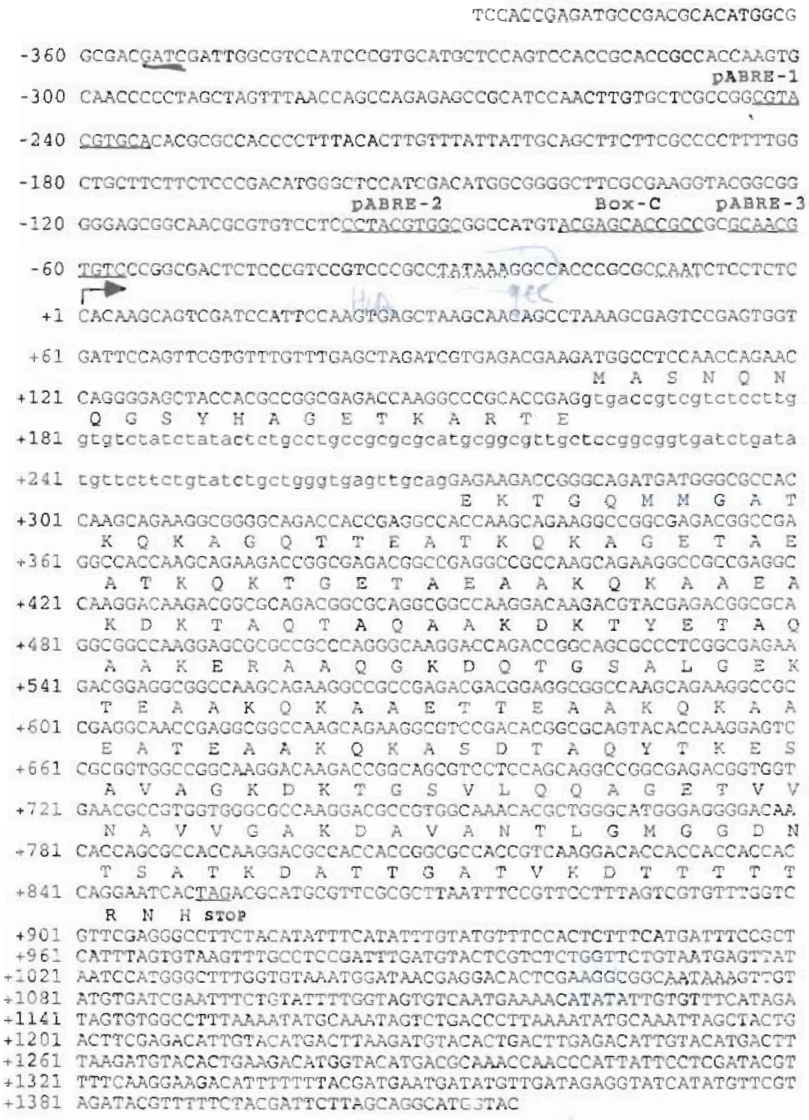

Fig. 4. Gene sequence and predicted amino acid sequence of *HVA1*. Nucleotides are numbered downstream from the transcription start site of + 1 under the arrow (see Fig. 5) or back from the same site in the upstream region. In the upstream region putative ABA responsive elements (pABRE) and a C-box are underlined and labeled. Elements of the basal transcription recognition complex, TATA and CAAT boxes have a dotted underline. The coding region is interrupted by a single intron, noted by lower-case nucleotides. The presumed polyadenylation signal is noted by a dotted underline in the 3'-untranslated region.

carrot *Dc8* and dicot soybean *GmPM2* which though similar to each other, were different enough from the rest of the sequences to be separated by a node on the tree with a significant bootstrap value of 85. The dicot *Brassica pLEA76*, monocot wheat *pMA1949*, and dicot cotton *D-29* were intermediate between the *HVA1-pMA2005* and *Dc3-D-7* clades and the distant carrot *Dc8* and soybean *GmPM2* groups. However, *Brassica pLEA76*, wheat *pMA1949*, and cotton *D-29* were different enough from each other to be placed on separate nodes by the consensus parsimony analysis but the bootstrap values were not significant (≤ 70) and did not support this conclusion. What the exact relationship of this intermediate group is could not be determined by the an[illegible]is

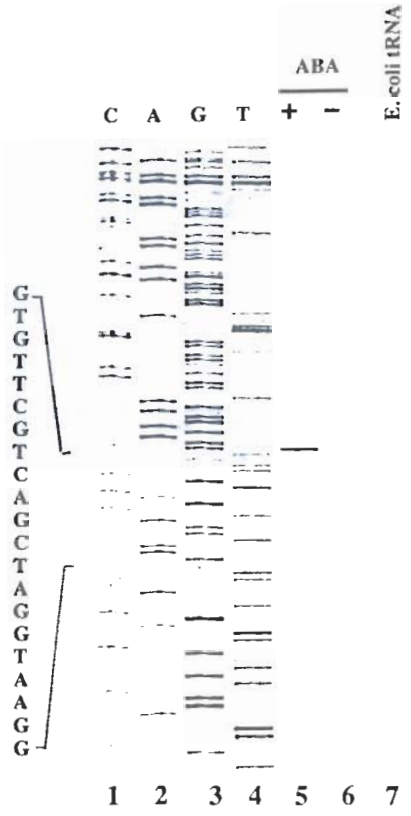

Fig. 5. Mapping of the HVA1 transcription start site by primer extension. Total RNA prepared from aleurone layers treated with or without 0.02 mM ABA for 24 h was used. The *E. coli* total RNA was included as a negative control. The size marker was obtained by performing sequencing reaction with the same oligonucleotide used for the primer extension as the primer and the HVA1 genomic clone as the template.

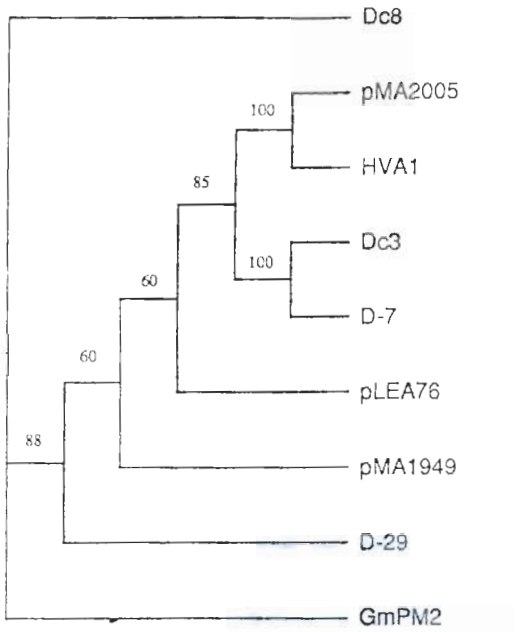

Fig. 6. Consensus phylogenetic tree constructed from *Lea 3* sequences: barley *HVA1*, wheat *pMA2005* and *pMA1949*, cotton *D-7* and *D-29*, carrot *Dc3* and *Dc8*, *Brassica pLEA76* and soybean *GmPM2*. Numbers on the nodes of the tree indicate bootstrap values with 1000 replicates from all possible tree topologies with a bootstrap value ≥70 indicating a ≥95% confidence level that the node represents a true phylogenetic inference.

but we report the tree as a tentative model of the infered evolution of the *Lea 3* gene family.

Transient expression assay for promoter function

To characterize the role of the putative ABA-responsive elements noted in the promoter region, the promoter, linked to a GUS reporter gene, was analyzed in a transient expression system utilizing biolistic transformation of barley aleurone layers. Since constructs containing the intron 1-exon 2-intron 2 fragment of the *HVA22* gene give the highest level of ABA response when linked to the GUS reporter and the *HVA22* promoter [42] all constructs used for this transient expression study contain the intron 1-exon 2-intron 2 fragment of the *HVA22* gene. The C-box sequence, which is found in several ABA-inducible genes, was mutated with the linker-scan technique [32] (Fig. 7). A mutation study on three sequences, pABRE1, pABRE2 and pABRE3 was also undertaken (Fig. 8), because these ABRE sequences are highly conserved in almost all ABA inducible genes, including *HVA22* [42], *Rab16* [21, 37], *Rab28* [38] and *Em* [21].

The wild type *HVA1* promoter responds to ABA, giving a 15-fold induction (WT, Fig. 7). When 9 bases of the C-box were mutated (LS-Box C, Fig. 7), no significant change was observed both at the transcription level and induction fold, suggesting that C-box is not involved in ABA response. Similarly, mutation of the whole pABRE1 sequence (CGTACGTGCA) had little effect on the response of the promoter to ABA

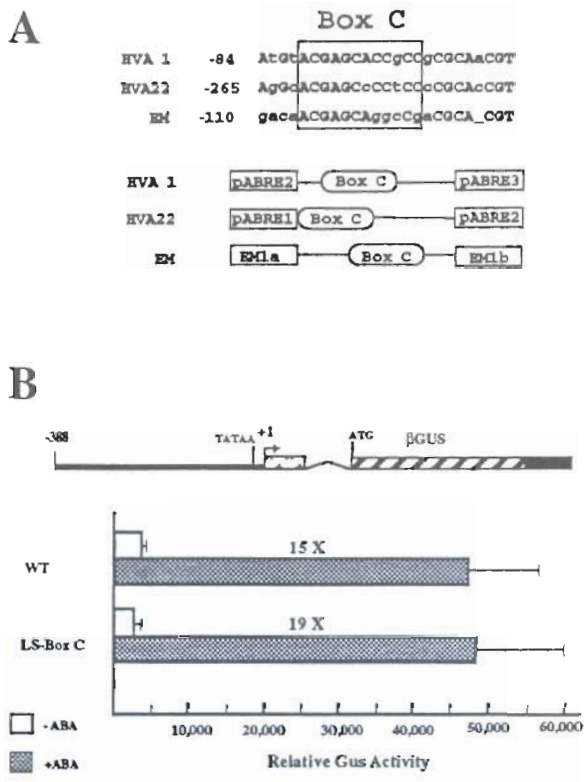

Fig. 7. Mutations in the conserved C box have no effect on the ABA induction of *HVA1*. A. The C-box sequence comparison is shown in the top panel and locations of C-box relative to other elements in the promoters of *HVA22*, *HVA1* and *Em* genes at the bottom. B. Analysis of C box mutations with the biolistic technique. A schematic diagram of testing construct is shown at the top; the thin black angled line indicates the position of the intron 1-exon 2-intron 2 fragment of HVA22 inserted between the 5'-untranslated sequence (thick black line) and the GUS-coding sequence (hatched box). The 3'-untranslated region (thick black line to the right of the hatched box) was from HVA22 genomic fragment including polyadenylation sequence (AATAAA). The white box stands for the relative GUS activity of the samples from layers incubated in the absence of ABA and the stippled box represents those from half seeds incubated in the presence of 0.02 mM ABA. The relative GUS activity, calculated as described in Materials and methods, of each construct is the mean of four replicas. Error bar indicates the standard error of each set of replicas. X indicates fold of increase.

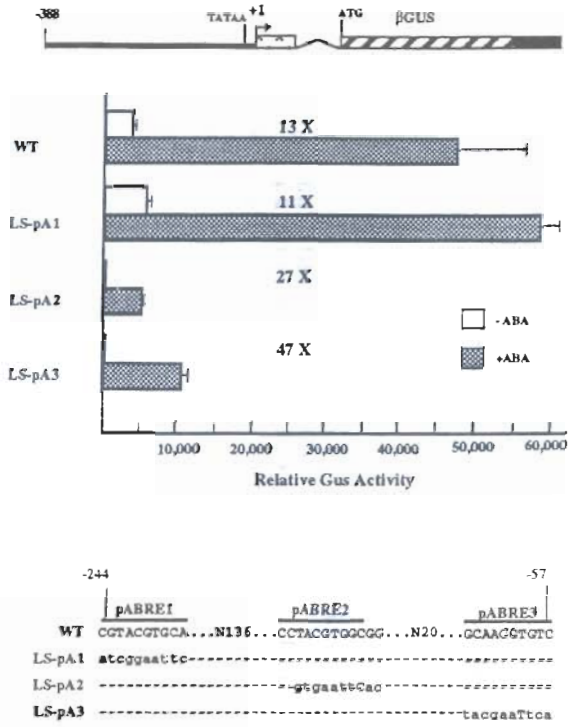

Fig. 8. Analysis of pABRE mutations in the *HVA1* promoter with the biolistic technique. The WT construct (wild type promoter) includes the sequence of HVA1 between −381 and +92. The mutated sequences in the LS-pA1, LS-pA2 and LS-pA3 constructs are shown at the bottom panel. All other labelings are the same as in Fig. 7.

(LS-pA1, Fig. 8). In contrast, when 9 out of 10 bases of the pABRE-2 sequence (TACGTG-GCGG), or the pABRE-3 sequence (GCAAC-GTGTC) were replaced, the level of transcription dropped to 10–20% of that obtained with the wild type promoter (LS-pA2, LS-pA3 and WT, Fig. 8). Interestingly, the mutated promoters containing mutated pABRE 2 to 3 gave a 27 or 47 induction fold response to ABA treatment, respectively even at the low level of transcription observed.

Discussion

The *HVA1* gene appears to be present as a single copy in the barley haploid genome, and its expression can be regulated by ABA and various stress conditions including drought, cold, heat and salinity. Using barley-wheat chromosome addition lines [30], we have tentatively mapped *HVA1* gene to barley chromosome 5 (B. Hong and T.D. Ho, unpublished). Based on sequence

comparisons, *HVA1* has been classified as a *Lea* gene [14]. In developing barley seeds, *HVA1* is mainly expressed in aleurone layers and peripheral regions of the embryo, yet no HVA1 transcripts or proteins are detected in the starchy endosperm [26]. We have estimated that the HVA1 protein constitutes about 1% of total soluble proteins in barley embryos. Since HVA1 is more highly expressed in the coleorrhiza and the epidermal layer of the coleoptile, the relative abundance of HVA1 in these tissues would be even higher than 1% of total protein [26]. Thus, the single-copy *HVA1* gene appears to be highly expressed in specific tissues of developing seeds. Despite being a *Lea* gene, the expression of HVA1 is certainly not limited to developing seeds as its expression can also be induced in vegetative tissues [27] (Fig. 1 of this work).

The barley *HVA1* gene was found to be structurally similar to the carrot *Dc3* [40] and the cotton *D-7* genes [2]. In these three genes the coding region starts with a short, less than 20 amino acids, first exon, interrupted by a short intron, followed by a larger downstream exon. Our parsimony sequence analysis indicates that the monocot barley *HVA1* and wheat *pMA2005* genes, that are highly similar, are most closely related to the dicot carrot *Dc3* and cotton *D-7* genes with which they share a common structural gene organization as just noted above. Interestingly, the other monocot and dicot sequences do not branch into two separate families indicating that divergence of the ancestral *Lea 3* genes probably occurred earlier than the divergence of the two families of the flowering plants. These results are consistent with the decision of Curry and Walker-Simmons [10] to postulate a subcategory of *Lea 3 (II)* because their wheat clone *pMA1949* is significantly different from their wheat clone *pMA2005*. The based their decision on similarities of *pM1949* to carrot *Dc8* and soybean *GmPM2*. However, it appears that wheat *pMA1949* is somewhat intermediate between the group related more closely to *HVA1* and the group related to carrot *Dc8*. The results of our parsimony analysis are encouraging. Construction of a better phylogeny awaits the identification of a *Lea 3* sequence from a more ancient group that could possibly serve as an ancestral model on which to base comparisons. Recently, proteins related to the group 2 Lea's, or dehydrins have been identified immunologically in the cyanobacteria [8]. As further sequences become available, phylogenetic inferences will become more reliable and a potential molecular clock may be used to calibrate the sequence evolution and make estimates on the time frames over which this divergence occurred.

The upstream promoter region of *HVA1* gene was found to contain sequence elements apparently conserved among ABA-inducible genes, three pABRE sequences [21, 37] and the C box which has not been reported before. However, mutations in either the pABRE1 region or the C box have no effect on the level of expression or the ABA inducibility of the *HVA1* gene. On the other hand, mutations of either pABRE2 (CCTACGTGGCGG) or pABRE3 (GCAACGTGTC), adversely affects the absolute level of gene expression. Surprisingly, these pABRE 2 & 3 mutants still display a high degree of ABA inducibility (Fig. 8). It is interesting to note that although all three pABREs in the *HVA1* promoter have the ACGT core that is conserved among all the ABREs, their role in ABA induction of *HVA1* appears to be quite different. Replacing all ten bp in pABRE1 including the ACGT core has no observable effect on the *HVA1* promoter. Thus, the presence of an ACGT core alone in a promoter region does not necessarily suggest a role in ABA response. In contrast, the presence of both pABRE 2 & 3 is necessary for high level of expression of *HVA1*, yet the ABA inducibility could be regulated by pABRE 2, 3, or another element. Sequences similar to the HVA1 pABRE have been shown to be essential for the ABA induction of several other genes, including the wheat *Em* [21], rice *Rab16* [44] and barley *HVA22* gene [42]. With the wheat *Em* gene, Gultinan *et al.* [21] have shown that a leucine-zipper protein, EmBP-1, binds to the pABRE sequence. A two base mutation of this sequence abolished the binding of EmBP-1 to the pABRE as well as the ABA responsiveness of the chimeric pro-

moter. The work of Skriver *et al.* [44] demonstrated that six copies of the sequence GTACGTGGCGC could confer ABA inducibility to a truncated 35S minimal promoter. In a similar study with another ABA-induced barley gene, *HVA22*, we have shown that the mutation of either the pABRE2 or pABRE3 sequence of *HVA22* reduces the expression and induction from this promoter in response to the ABA treatment [42]. However, the mutated promoter still responds to the ABA treatment at a lower level. The double mutation of pABRE2 and pABRE3, in contrast, completely abolishes the ABA induction of the *HVA22* promoter [41]. Taking all these data together, it is apparent that the so-called ABREs have not yet been clearly defined. As pointed out by Guiltinan *et al.* [21], many ABA-insensitive gene promoters also contain sequences similar to ABREs. Even the presence of a highly conserved ACGT core does not necessarily confer ABA response as shown in the case of barley *HVA1* pABRE 1. Other surrounding sequences, which are not yet defined, must play a role in conferring ABA response and differentiating this response from other developmental and environmental responses linked to G box elements containing ACGT cores. Towards this end, Williams *et al.* [48] systematically replaced the border nucleotide sequences surrounding the consensus ACGT core in a series of oligonucleotides and looked at binding affinities of these oligonucleotides with nuclear binding proteins. This study indicated that the nucleotide sequences surrounding the ACGT core defined the affinity for a particular type of nuclear binding protein and the strength of binding observed. Thus it is conceivable that the flanking sequences of the ACGT core could determine the specific function of a G box.

Our analyses of the barley *HVA1* gene presented in this paper also strongly suggest that, at least in the context of native promoters, multiple regulatory elements need to work together in order to confer a high level of ABA-regulated gene expression. A similar case has been shown in our recent studies with the barley *HVA22* gene, where two elements in the promoter and intron 1 have to be present to have the full ABA response [41]. In another hormone responsive gene, the barley amylase gene, Lanahan *et al.* [33] have reported that as many as four promoter elements have to interact with one another in order to confer gibberellic acid induction of this gene in the aleurone tissue during post-germination growth. In mammalian systems, the composite glucocorticoid response element appears to have subsites capable of interacting with both the glucocorticoid receptor protein and another transcription factor, AP-1 [12]. In summary, the ABA induction of barley *HVA1* gene appears to be regulated by a multi-element complex: pABRE 2 & 3 working together to regulate the absolute level of expression; and either of these elements or a possible third element may regulate ABA inducibility. Work is currently underway to define this last element as well as to investigate the interactions among these elements.

Acknowledgements

We thank J. Han for her technical assistance and T.-Y. Chiang for his assistance in performing and interpreting the phylogenetic analysis. We are grateful for Bimei Hong for performing the experiments presented in Fig. 1. This research was funded by grants from the USDA (No. 910302), Consortium for Plant Biotechnology Research and Ciba-Geigy Inc. to T.H.D. Ho.

References

1. Anderberg RJ, Walker-Simmons MK: Isolation of a wheat cDNA clone for an abscisic acid-inducible transcript with homology to protein kinase. Proc Natl Acad Sci USA 89: 10183–10187 (1992).
2. Baker J, Steele C, Dure-III L: Sequence and characterization of 6 *Lea* proteins and their genes from cotton. Plant Mol Biol 11: 277–291 (1988).
3. Bartels D, Engelhardt K, Roncarati R, Schneider K, Rotter M, Salamini F: An ABA and GA modulated gene expressed in the barley embryo encodes an aldose reductase related protein. EMBO J 10: 1037–1043 (1991).
4. Belanger FC, Brodl MR, Ho T-HD: Heat shock causes destabilization of specific mRNAs and destruction of en-

doplasmic reticulum in barley aleurone cells. Proc Natl Acad Sci USA 83: 1354–1358 (1986).
5. Brown J: A catalogue of splice junction and putative branch point sequences from plant introns. Nucl Acids Res 14: 9549–9559 (1986).
6. Bruce WB, Quail PH: *Cis*-acting elements involved in photoregulation of an oat phytochrome promoter in rice. Plant Cell 2: 1081–1089 (1990).
7. Church GM, Gilbert W: Genomic sequencing. Proc Natl Acad Sci USA 81: 1991–1995 (1984).
8. Close TJ, Lammers PJ: An osmotic stress protein of cyanobacteria is immunologically related to plant dehydrins. Plant Physiol 101: 773–779 (1993).
9. Curry J, Morris CF, Walker-Simmons MK: Sequence analysis of a cDNA encoding a group 3 LEA mRNA inducible by ABA or dehydration stress in wheat. Plant Mol Biol 16: 1073–1076 (1991).
10. Curry J, Walker-Simmons MK: Unusual sequence of group 3 LEA (II) mRNA inducible by dehydration stress in wheat. Plant Mol Biol 21: 907–912 (1993).
11. Devereux J, Haeberli P, Smithies O: A comprehensive set of sequence analysis programs for the VAX. Nucl Acids Res 12: 387–395 (1984).
12. Diamond MI, Miner JN, Yoshinaga SK, Yamamoto KR: Transcription factor interactions: selectors of positive or negative regulation from a single DNA element. Science 249: 1266–1272 (1990).
13. Dure-III L: A repeating 11-mer amino acid motif and plant desiccation. Plant J 3: 363–369 (1993).
14. Dure-III L, Crouch M, Harada J, Ho THD, Mundy J, Quatrano R, Thomas T, Sung ZR: Common amino acid sequence domains among LEA proteins of higher plants. Plant Mol Biol 12: 475–486 (1989).
15. Dure-III L, Greenway SC, Galau GA: Developmental biochemistry of cottonseed embryogenesis and germination. XIV. Changing mRNA populations as shown by *in vitro* and *in vivo* protein synthesis. J Biochem 20: 4162–4168 (1981).
16. Efron B: The jacknife, the bootstrap and other resampling plants. Conf Board Math Sci Soc Ind Appl Math 38: 1–92 (1982).
17. Feinberg AP, Vogelstein B: A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132: 6–13 (1983).
18. Felsenstein J: Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39: 783–791 (1985).
19. Franz G, Hatzopoulos P, Jones TJ, Krauss M, Zung ZR: Molecular and genetic analysis of an embryonic gene, DC 8, from *Daucus carota* [L.]. Mol Gen Genet 218: 143–151 (1989).
20. Gilman M: Ribonuclease protection assay. In: Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K (eds) Current Protocols in Molecular Biology, pp. 4.7.1–4.7.8. John Wiley, New York (1989).
21. Guiltinan MJ, Marcotte WR Jr., Quatrano RS: A plant leucine zipper protein that recognizes an abscisic acid response element. Science 250: 267–271 (1990).
22. 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 Mol Biol 12: 395–401 (1989).
23. Henikoff S: Unidirectional digestion with exonuclease III creates targetted breakpoints for DNA sequencing. Gene 28: 351–359 (1984).
24. Hillis DM, Bull JJ: An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Syst Biol 42: 182–192 (1993).
25. Hillis DM, Huelsenbeck JP: Signal, Noise, and reliability in molecular phylogenetic analyses. J Hered 83: 189–195 (1992).
26. Horig B: Regulation of synthesis and potential function of an ABA- and stress-induced protein in barley. Ph.D. thesis, Washington University, St Louis, MO (1991).
27. Hong B, Barg R, Ho THD: Developmental and organ-specific expression of an ABA- and stress-induced protein in barley. Plant Mol Biol 18: 663–674 (1992).
28. Hong B, Uknes SJ, Ho THD: Cloning and characterization of a cDNA encoding a mRNA rapidly-induced by ABA in barley aleurone layers. Plant Mol Biol 11: 495–506 (1988).
29. Hsing YC, Chen ZY, Chow TY: Nucleotide sequences of a soybean complementary DNA encoding a 50-kilodalton late embryogenesis abundant protein. Plant Physiol 99: 353–355 (1992).
30. Islam AKMR, Shepherd KW, Sparrow DHB: Isolation and characterization of euplasmic wheat-barley chromosome addition lines. Heredity 46: 161–174 (1981).
31. Jefferson RA, Kavanagh TA, Bevan MW: GUS fusions: β-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J 6: 3901–3907 (1987).
32. Kunkel TA, Roberts JD, Zakour RA: Rapid and efficient site-specific mutagenesis without phenotypic selection. Meth Enzymol 154: 367–382 (1987).
33. Lanahan MB, Ho T-HD, Rogers SW, Rogers JC: A gibberellin response complex in cereal a-amylase gene promoters. Plant Cell 4: 203–211 (1992).
34. Ludevid MD, Freire MA, Gomez J, Burd CG, Albericio F, Giralt E, Dreyfuss G, Pages M: RNA binding characteristics of a 16 kDa glycine-rich protein from maize. Plant J 2: 999–1003 (1992).
35. Maniatis T, Fritsch EF, Sambrook J: Molecular Cloning: A Laboratory Manual, vols. 1–3. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989).
36. Michel D, Salamini F, Bartels D, Dale P, Baga M, Szalay A: Analysis of a desiccation and ABA-responsive promoter isolated from the resurrection plant *Craterostigma plantagineum*. Plant J 4: 29–40 (1993).
37. Mundy J, Yamaguchi-Shinozaki K, Chua N-H: Nuclear proteins bind conserved elements in the abscisic acid-responsive promoter of a rice *rab* gene. Proc Natl Acad Sci USA 87: 1406–1410 (1990).

38. Pla M, Vilardell J, Guiltinan MJ, Marcotte WR, Niogret MF, Quatrano RS, Pages M: The *cis*-regulatory element CCACGTGG is involved in ABA and water-stress responses of the maize gene *rab28*. Plant Mol Biol 21: 259–266 (1993).
39. Sanger F, Nicklen S, Coulson AR: DNA sequencing with chain terminating inhibitors. Proc Natl Acad Sci USA 74: 5463–5467 (1977).
40. Seffens WS, Almoguera C, Wilde HD, Vonder-Haar RA, Thomas TL: Molecular analysis of a phylogenetically conserved carrot gene: developmental and environmental regulation. Devel Genet 11: 65–76 (1990).
41. Shen Q: Molecular mechanism of gene regulation by the phytohormone abscisic acid. Ph.D. thesis, Washington University, St Louis, MO (1993).
42. Shen Q, Uknes SJ, Ho THD: Hormone response complex of a novel abscisic acid and cycloheximide inducible barley gene. J Biol Chem 268: 23652–23660 (1993).
43. Skriver K, Mundy J: Gene expression in response to abscisic acid and osmotic stress. Plant Cell 2: 503–512 (1990).
44. Skriver K, Olsen FL, Rogers JC, Mundy J: *Cis*-acting DNA elements responsive to gibberellin and its antagonist abscisic acid. Proc Natl Acad Sci USA 88: 7266–7270 (1991).
45. Sutton WD: Some features of the DNA of *Rhizobium* bacteroids and bacteria. Biochim Biophys Acta 366: 1–10 (1974).
46. Swoford DL: PAUP: Phylogenetic Analysis Using Parsimony, Ver. 3.0. Illinois Natural History Survey, Champaign, IL (1990).
47. Urao T, Yamaguchi-Shinozaki K, Urao S, Shinozaki K: An Arabidopsis *myb* homolog is induced by dehydration stress and its gene products binds to the conserved MYB recognition sequence. Plant Cell 5: 1529–1539 (1993).
48. Williams ME, Foster R, Chua N-H: Sequences flanking the hexameric G-box core CACGTG affect the specificity of protein binding. Plant Cell 4: 485–496 (1992).