Molecular cloning and characterization of a group 3 LEA gene from *Agropyron mongolicum* Keng

Yan Zhao¹, Jinfeng Yun¹*, Fengmin Shi¹, Junjie Wang¹, Qingchuan Yang² and Yuehui Chao³

¹College of Ecology and Environmental Science, Inner Mongolia Agricultural University, Hohhot, Inner Mongolia 010018, R.P of China.
²Institute of Animal Sciences, Chinese Academy of Agricultural Sciences, Beijing 100193, China.
³Bioengineering College of Chongqing University, Chongqing 400030, People’s Republic of China.

Accepted 5 August, 2010

Late embryogenesis-abundant (LEA) protein is one of the components involved in desiccation tolerance (DT) by maintaining cellular structures in the dry state. In this study, a member of the group 3 LEA, MwLEA1, was cloned from Mongolian wheatgrass (*Agropyron mongolicum* Keng) based on a homologous sequence from wheat (*Triticum aestivum* L.). Its full-length cDNA sequence was 705 bp, encoding a protein of 187 amino acids. The amino acid sequence comparison revealed its high homology with LEA proteins from other plant species. The deduced MwLEA1 protein had five repeat 11-amino-acid motifs, with a molecular weight of 19.4 kDa and a theoretical isoelectric point of 8.8. Subcellular localization indicated that the MwLEA1 was localized in the nucleus of the onion epithelial cell. Under water stress conditions, MwLEA1 exhibited different expression levels, which was higher in root and shoot but lowest in leaf. The expression profiling under different stresses indicated that MwLEA1 played roles in responses to water, salt stresses as well as abscisic acid (ABA) regulation. The gene of MwLEA1 was transformed into tobaccos by *Agrobacterium tumefaciens*-mediated method. Eleven regenerated plants were analyzed by polymerase chain reaction (PCR) and southern blotting, and 6 of them were proved to be transgenic plants.

Key words: *Agropyron mongolicum* Keng, cloning, late embryogenesis abundant, subcellular localization, expression, transformation.

INTRODUCTION

Late embryogenesis abundant (LEA) proteins could accumulate to high levels during the last stage of seed maturation and during water deficit in vegetative organs, suggesting a protective role during water limitation (Dure, 1993b; Bray, 1997; Garay-Arroyo et al., 2000; Hoekstra et al., 2001). LEA proteins were first isolated in cotton as a set of proteins (Dure et al., 1981). Subsequently, many LEA proteins or their genes have been characterized from different plant species (Dure, 1992).

LEA proteins are classified according to the appearance of different sequence motifs/patterns or biased amino acid composition; plant LEA proteins have been separated into different groups (Bray, 1993; Dure, 1989; Dure, 1993; Wise, 2003). More common classifications are related to their protein structural domains or chemical characteristics. Currently, two criteria are used for the classification of LEA proteins. The traditional one was first described by Dure et al. (1989). As further studies were carried out in plants and more recently other type of organism, including invertebrates and microorganisms (Browne et al., 2004; Tunnacliffe et al., 2007).

Among the LEA protein members, Group 3 LEA was characterized as containing a repeat of an 11-amino acid...
motif (TAQAAKEKAXE) that may form an amphiphilic α-helix structure (Baker et al., 1988; Dure et al., 1989; Dure, 1993; Tunnacliffe et al., 2007). These genes have been reported in several plant species. Lea76 of *Brassica napus* contains thirteen repeats of a homologous amino acid motifs (Harada et al., 1989), and HVA1 of barley contains nine repeats (Hong et al., 1988). Group 3 LEA proteins with similar repeating amino acid motifs have also been reported in cotton (Baker et al., 1988), carrot (Choi et al., 1987; Dure et al., 1989) and rapeseed (Harada et al., 1989). The presence of group 3 LEA mRNAs in these plant species indicates that there may be a common mechanism for desiccation protection. The presences of amino acid residue distribution along the polypeptides, the predicted novel structure of the LEA proteins and the correlation of LEA gene expression with physiological and environmental stresses in the transgenic plants provide evidences that LEA protein may play a protective role in plant cells under various stress conditions; moreover, this protective role may be essential for the survival of the plant under extreme stress conditions (Baker et al., 1988; Dure et al., 1989; Skriver and Mundy, 1990; Chandler and Robertson, 1994). However, direct experimental evidence supporting the exact functions of LEA proteins is still lacking and the physiological roles of LEA proteins remain largely unknown.

Mongolia wheatgrass (*Agropyron mongolium* Keng) (2n = 2x = 14), an important wild relative species of wheat, is a perennial, cross-pollination diploid, which is distributed in Desert Grassland and Typical Grassland of China. Long-term evolution and adaptation to harsh conditions make Mongolia wheatgrass rich in tolerance genes for a range of biotic and abiotic stresses such as pest and fungal attacks, drought, cold, barren and high salinity (Yun et al, 1989; Yu et al., 2002; Qi et al, 1998). In view of all these attributes, Mongolia wheatgrass has been proposed to be a valuable genetic resource in forage grass and crop improvement for resistances or tolerances. There are abundant germplasm resources of Mongolia wheatgrass in Northwest China. However, up till now, none of the LEA genes has been isolated and characterized from Mongolia wheatgrass. In the present study, the first full-length LEA protein cDNA, designated as *MwLEA1*, from Mongolia wheatgrass was obtained by reverse transcriptase polymerase chain reaction (RT-PCR). The gene was successful transformed into tobaccos. The cloning of the gene makes it possible to enhance the abiotic stress tolerances by genetic engineering in forage grasses. The characterization of *MwLEA1* might also provide further insight into the physiological processes of stress response in higher plants.

**MATERIALS AND METHODS**

**Plant materials and growth conditions**

Seeds of wild Mongolia wheatgrass species were collected from its natural habitat in XilinGol prairier, Inner Mongolian, China. The materials were grown in a greenhouse or experimental field under a normal day-length period. After 4 - 8 weeks, the seedlings were treated with 0.1mM/L abscisic acid (ABA), 0.25 mM/L NaCl, 20% PEG6000 or drought for 7 days, and were then harvested and frozen in liquid nitrogen and stored at -80°C until use.

**Amplification of full-length *MwLEA1* cDNA**

Total RNA was isolated from Mongolia wheatgrass seedlings of 35-day-old using a guanidine thiocyanate extract method (Cathala et al., 1983). Any contaminated genomic DNA was removed by incubating the total RNA with RNase-free DNase (Promega) at 37°C for 30 min. Total RNA was used to synthesize cDNA, and cDNA was used as template for further amplification. Based on the LEA protein sequence of wheat from GenBank (Accession number: AY148492), gene-specific primers: kf (5' - ATGGCCTCCAACCGAAACCCA -3') and Kr (5' - CGAGGAGTACATCAAATGGA -3') were synthesized and used to amplify the target gene using the Advantage™ 2 PCR Enzyme Kit (Clontech). PCR was performed with 30 cycles at 94°C for 30 s, 58°C for 30 s and 72°C for 1 min, and an additional poly-merization step at 72°C for 10 min. The PCR product was separated by electrophoresis on a 1% agarose gel stained with ethidium bromide, and purified using the DNA Gel Extraction Kit (TaKaRa, Japan). The purified products were cloned into the pMD-19T (TaKaRa, Japan) vector and then transformed into Escherichia coli DH5α. The recombinant plasmids were sequenced by Shanghai Sangon Biological Engineering Technology & Services (Shanghai, China).

**Sequence analysis**

Sequences were analyzed with DNAman and DNAuser soft. Sequence homology analysis was against nucleotide and protein database of GenBank using, Basic Local Alignment Search Tool (BLAST) tools. Conserved sequences were analyzed with the program of Conserved Domains (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) and ExPaSy ScanProsite (http://www.expasy.org/tools/scanprosite/). Potential signal peptide cleavage site was identified using Signal P 3.0 (http://www.cbs.dtu.dk/services/SignalP/); and pl/MW was predicted on the ExPaSy website (http://us.expasy.org/tools/pi_tool.html).

**Subcellular localization of *MwLEA1***

*MwLEA1*-GFP (green fluorescent protein) fusion vector was constructed and transformed into onion epidermal cells to express the fusion protein. To amplify the coding sequence of *MwLEA1*, two primers were designed, one with the XhoI restriction site (Yf: 5' - ctcgac ATGGGCCATCCAGGACCCGAACT -3') and another with the SalI restriction site (Yr: 5' - ggtcga GTGAATTCGGTGATGCTGCTTTTCT -3'). The PCR product was digested with XhoI and SalI, and ligated with the vector pAT-GFP which was digested with the same restriction enzymes. The fusion vector was then transformed into onion epidermal cells using a gene gun. Subcellular localization of transiently expressed *MwLEA1*-GFP fusion protein was detected by a confocal laser scanning microscope (Nikon, A1Rsi Confocal).

**Expression analysis of *MwLEA1***

RT-PCR was performed using primers which were previously used for amplification of *MwLEA1*, kf and kr. Young leaves, stems and roots were collected from 35-day-old seedlings. The 35-day-old seedlings were treated with 0.25 mM NaCl, 20% PEG6000, 0.1 mM ABA for 2, 4, 6, 8 and 10 h, and the seedlings without treatment
Figure 1. The nucleotide sequence of MwLEA1 cDNA and the deduced amino acid sequence. The amino acids shaded in black are 11-amino acids motifs. Total amino acid number: 187; MW, 19241; (1-565); Max ORF starts at AA pos 1 for 187 AA (561 bases).

were used as control. The seedlings were treated drought to 7 days for tissue specific expression. Total RNA of each was reverse transcribed and the synthesized cDNA was used as template for RT-PCR.

Transformation of MwLEA1 into tobacco

Tobacco leaf discs (5 – 10 mm in length) were infected with Agrobacterium tumefaciens LBA4404 strain containing pBI::MwLEA1. After 3 days of co-culture in the dark at 28°C, the leaf discs were transferred to the regeneration medium supplemented with Rifampicin (50 µg /ml) and kanamycin (50 µg /ml). The cultures were maintained at 28°C under continuous illumination. Shoot bud differentiation started after 14 – 16 days of culture, which elongated into shoots within 30 – 35 days. After profuse rooting in the rooting medium, the T0 plants were identified by PCR and positive plants were allowed to flower and seeds were set under normal environmental conditions.

PCR and Southern blot

Genomic DNA was extracted from the young leaf tissues of transgenic tobacco plants according to the modified cetyl trimethylammonium bromide (CTAB) method (Murray and Thompson, 1980; Zhang et al., 1992). The gene-specific primers (5’-ATGGCC TCAAACGAAACC-3’ and 5’-GTGATTGCCGTTGATCTTC -3’) were used to identify positive transgenic plants. PCR reaction was conducted in a volume of 20 µl containing 100 ng genomic DNA, 2 mM MgCl2, 0.2 mM of each dNTP, 1×PCR buffer, 0.2 mM of each primer and 1 unit rTaq polymerase (TaKaRa). The PCR reaction was performed at 94°C for 5 min; then with 30 cycles of 94°C for 30 sec, 60°C for 30 sec, 72°C for 1 min and finally at 72°C for 10 min. For PCR-Southern blot hybridization, 6 of each of the transgenic plants samples were transferred onto the OPTITRAN pure nitrocellulose membrane (Schleicher and Schuell, USA) using 20×SSC solution as transfer buffer (Koetsier et al., 1993). The 587 bp fragment of MwLEA1 gene was DIG-labeled using the DIG High Primer DNA Labeling and Detection Starter Kit I (Roche). Hybridization was performed at 42°C overnight in the DIG Easy Hyb and detection was performed at high stringency conditions following the standard protocol (Sambrook and Russell, 2001).

RESULTS

Molecular cloning and sequence analysis of MwLEA1

The full-length cDNA of target gene isolated from Mongolia wheatgrass was 705 bp, which was designated as MwLEA1 (GenBank accession no: GU724972). Sequence analysis of MwLEA1 indicated that its full-length cDNA contained an open reading frame (ORF) of 565 bp encoding a putative protein of 187 amino acids. After the stop code TAG, there was a 160-bp 3'-UTR. The putative protein had a conserved sequence of five 11-amino acids motifs (Figure 1). Its amino acid sequence was used to search the protein databank and the result showed that the MwLEA1 could be aligned with the other group 3 LEA proteins from different species (Figure 2). The theoretical
The DNAman soft predicted that the secondary structure of MwLEA1 protein was mainly α-helices structure and forming hydrophilic climax. The amino acid sequence composition was 14.08% alanine (Ala), 15.40% lysine (Lys), 13.59% threonine (Thr) and without any cysteine (Cys), tryptophan (Trp) and proline (Pro). No signal peptide was found using the signal peptide analysis.

**Figure 2.** Comparison of amino acid sequences of MwLEA1 with other reported zinc finger proteins in the National Center for Biotechnology Information (NCBI) database. Aad02421 is group 3 LEA protein in Oryza sativa; bac80266 is ABA inducible protein in Triticum aestivum; baf7927 is group 3 late embryogenesis abundant protein in Triticum aestivum; cca55041 is HVA1 in Hordeum vulgare subsp. Vulgare.

**Subcellular localization of MwLEA1**

The MwLEA1-GFP fusion vector was constructed to examine the subcellular localization of MwLEA1. Both the fusion and pA7-GFP vectors were transformed into onion epidermal cells by a gene gun. The cultured onion epidermal cells transformed with MwLEA1-GFP fusion vector showed a strong fluorescence signal only in the...
nucleolus under the confocal laser scanning microscope (Figure 3C and D), while the GFP signal distributed throughout the onion epidermal cells when transformed with the control pA7-GFP vector (Figure 3A and B). This result indicated that MwLEA1 was a nuclear-localized protein.

**Expression analysis of the MwLEA1 gene**

To examine the expression pattern of the MwLEA1 gene in Mongolia wheatgrass, cDNAs from various Mongolia wheatgrass tissues and Mongolia wheatgrass seedlings induced by 0.25 mM NaCl, 20% PEG6000 and 0.1 mM ABA for different times were used as templates for RT-PCR analysis using gene-specific primers. Housekeeping gene, the actin gene of Mongolia wheatgrass was used to normalize the target gene. Endogenous MwLEA1 gene was found in all organs tested, including root, stem and leaf. The expression levels in root and stem were higher (Figure 4), and that in leaf was the lowest. The expression of MwLEA1 could be induced by both 20% PEG6000 and 0.1 mM ABA at different times. However, their expression pattern was different. The expression reached peak at 6 h after 20% PEG6000 (Figure 5) treatment and 8 h after 0.1 mM ABA treatment (Figure 6). The expression of MwLEA1 was found to be slightly induced by 0.25 mM NaCl, and remained at a stable relatively low level (Figure 7). These results imply that MwLEA1 might play roles in the responses of Mongolia wheatgrass to water stress, salt stress and ABA regulation.

**Transformation of MwLEA1 gene into tobacco and characterization of the transgenic plants**

The plant expressing vector pBl: MwLEA1 was constructed with the full length ORF of MwLEA1 gene and expression vector pBl121, and then transformed into tobacco by A. tumefaciens LBA4404-mediated method. About one hundred regenerated plants were produced (Figure 8), and 11 regenerated plants were analyzed by PCR and PCR-Southern blotting analysis with wild type tobacco and the pBl: MwLEA1 as control. Six of them...
Figure 4. RT-PCR amplification of MwLEA1 in different organs.

Figure 5. Expression patterns of MwLEA1 under 20%PEG treatments. Effect of abiotic treatments on MwLEA1 expression was determined by RT-PCR. The numbers above each lane indicate the time in hours (h) after the initiation of treatment. All RT-PCR expression assays were performed and analyzed at least three times in independent experiments. Only the optimal pictures were shown. M, Molecular mass marker DL1500; 3 and 10 are expression of leaves, and others are expression of roots-stems.

Figure 6. Expression patterns of MwLEA1 under 0.1mmol/L ABA treatments. Effect of abiotic treatments on MwLEA1 expression was determined by RT-PCR. The numbers above each lane indicate the time in hours (h) after the initiation of treatment. All RT-PCR expression assays were performed and analyzed at least three times in independent experiments. Only the optimal pictures were shown. M, Molecular mass marker DL1500; 1 and 8 are expression of leaves, and others are expression of roots-stems.

Figure 7. Expression patterns of MwLEA1 under 0.25 mol/L NaCl treatments. Effect of abiotic treatments on MwLEA1 expression was determined by RT-PCR. The numbers above each lane indicate the time in hours (h) after the initiation of treatment. All RT-PCR expression assays were performed and analyzed at least three times in independent experiments. Only the optimal pictures were showed. M, Molecular mass marker DL1500.
proved to be positive transgenic plants (Figure 9). The successful transformation of MwLEA1 gene into tobacco provided useful materials for further characterization of the functions of this gene.

DISCUSSION

In general, LEA proteins are relatively small, with most falling in a range from 10 to 30 kDa (Michaela and Dirk, 2008). Most LEA peptides are reported highly hydrophilic and heat soluble and contain some amino acid residues in abundance (e.g., Gly, Ala, and Glu) and others in scarcity (e.g., Cys and Trp) (Dure et al., 1989; Dure, 1993). TAQAKÉKAGE of the 11-mer amino acid motifs, is a typical structure of group 3 LEA protein. Lea76 of B. napus contains thirteen repeats of a homologous amino acid tract and HVA1 of barley contains nine repeats. Group 3 LEA proteins with similar repeating amino acid tracts have also been reported in cotton, carrot, and rapeseed. It may form an amphiphilic a-helix structure (Baker et al., 1988; Dure et al., 1989; Dure, 1993). The arrangement of charged amino acids within the motif suggested a function in sequestering ions that will accumulate under dehydration conditions (Dure, 1993). In this study, we isolated the full-length cDNA of MwLEA1 from Mongolia wheatgrass. Sequence analysis of MwLEA1 indicated that it contained five conserved 11-mer amino acid motifs, and showed high identity with group 3 LEA proteins, especially with group 3 LEA proteins of wheat (Figures 1 and 2). The α-helices structure was the main secondary structure of MwLEA1, and formed hydrophilic climax predicted by DNAman soft. Ala, Lys and Thr were about 50% in amino acid composition of sequence, without Cys, Trp and Pro. Together, these findings suggest that the MwLEA1 might be a member of group 3 LEA proteins and play important roles in modulating the tolerance to water stress. However, the exact role of this gene is still unknown.

LEA proteins are also reported to be present in abundance in seeds during the late stages of development, and they are associated with desiccation tolerance (Baker et al., 1988; Dure et al., 1989; Blackman et al., 1991), ABA content (Galau et al., 1986, 1992) or subsequent seed germination (Ried and Walker-Simmons, 1993). In addition, most LEA proteins accumulate in immature seeds, seedlings or vegetative tissues with exogenous ABA, drought, salinity or temperature stress (Dure et al., 1989; Wang et al., 2003). Water stress is one of the major adverse environmental conditions that affect plant growth, development and crop yield. So far, researches on the LEA protein have mainly focused on Arabidopsis thaliana and rice, and no LEA protein in the water stress response pathway in Mongolian wheatgrass has been identified.

We have demonstrated that the transcriptional expression of the MwLEA1 gene was quickly and transiently induced by PEG6000, exogenous ABA and salt stress. These results implied that the MwLEA1 might be involved...
in water, salt and ABA inducing responses in Mongolian wheatgrass. Previous studies have demonstrated that some LEA proteins played important regulatory roles in different abiological processes by modulating the mRNA processing or degradation of key components in various organisms. The identification of the MwLEA1 will help to understand how these groups of gene act in the regulation of water, salt stresses and ABA inducing response. In response to dehydration, ABA levels increased dramatically and group 3 LEA mRNAs were induced in root, shoot and scutellar tissue. However, it is also reported that group 3 LEA proteins were detected only in shoot and scutellar tissue and not in root tissue of wheat seedlings. Treatment of unstressed seedlings with 20 PM ABA resulted in low levels of group 3 LEA proteins in the roots, whereas, higher levels were found in the shoot and scutellar tissue (Jeffrey and Walker-Simmons, 1993). Under water stress, accumulations of high level of MwLEA1 were observed in the root and shoot. Probably, one reason is that Mongolian wheatgrass has the sandy coat for rooting system (Gao et al., 1990) and a little share of leaf. It mainly depends on the rooting system as well as keeping the growth and development under severely dehydrated conditions. Thus, in Mongolian wheatgrass seedlings severely dehydrated, the accumulation of high levels of group 3 LEA proteins is correlated with tissue dehydration tolerance.

Subcellular localization analysis indicated that MwLEA1 is localized in the nucleus. Many LEA proteins were reported to be localized in the cytoplasm. TaLEA1 in wheat and Dc8 in carrot are localized in the cytoplasm (Gerald et al., 1989), protein bodies and cell walls of zygotic embryo and endosperm tissues. Others, such as RAB28 in Arabidopsis, were reported to be localized in the nucleus. From both prediction and published experimental evidences, the LEA proteins can be present in all subcellular compartments. Whether they have different functions in different compartments and what function they possess need to be further determined.

In conclusion, a MwLEA1 gene, encoding a water stress-induced LEA protein, was isolated from Mongolian wheatgrass for the first time. The putative protein was localized mainly in the nucleus, and had no signal peptide. There are five conserved 11-mer amino acid motifs. The gene might have function in drought and salt tolerances. MwLEA1 genes were successfully transformed into tobacco, and the transgenic plants will be useful materials for the characterization of its function.

ACKNOWLEDGEMENTS

We thank Prof. Xiue Wang for advice and help and Prof. Qingchuan Yang for the generous gifts of green fluorescent protein vector. This work was supported by grants from the National Basic Research Program of China (973 Program No. 2007CB108901-1) and the National Nature Science Foundation of China (No. 30760159).

REFERENCE


