Appl Biochem Biotechnol (2009) 157:244–253
DOI 10.1007/s12010-008-8398-6

Isolation and Characterization of an Atypical LEA Protein Coding cDNA and its Promoter from Drought-Tolerant Plant *Prosopis juliflora*

Suja George · B. Usha · Ajay Parida

Received: 14 February 2008 / Accepted: 9 October 2008 /
Published online: 12 November 2008

Abstract Plant growth and productivity are adversely affected by various abiotic and biotic stress factors. Despite the wealth of information on abiotic stress and stress tolerance in plants, many aspects still remain unclear. *Prosopis juliflora* is a hardy plant reported to be tolerant to drought, salinity, extremes of soil pH, and heavy metal stress. In this paper, we report the isolation and characterization of the complementary DNA clone for an atypical late embryogenesis abundant (LEA) protein (Pj LEA3) and its putative promoter sequence from *P. juliflora*. Unlike typical LEA proteins, rich in glycine, Pj LEA3 has alanine as the most abundant amino acid followed by serine and shows an average negative hydropathy. Pj LEA3 is significantly different from other LEA proteins in the NCBI database and shows high similarity to indole-3 acetic-acid-induced protein ARG2 from *Vigna radiata*. Northern analysis for Pj LEA3 in *P. juliflora* leaves under 90 mM H_2O_2 stress revealed up-regulation of transcript at 24 and 48 h. A 1.5-kb fragment upstream the 5′ UTR of this gene (putative promoter) was isolated and analyzed in silico. The possible reasons for changes in gene expression during stress in relation to the host plant's stress tolerance mechanisms are discussed.

Keywords *Prosopis juliflora* · Abiotic stress · Atypical LEA proteins · Promoter · Phylogenetic analysis

Introduction

Plant growth and productivity are adversely affected by various abiotic and biotic stress factors. Abiotic stress is the principal cause of crop failure worldwide, causing a dip in average yields for most major crops by more than 50% [1, 2]. The development of crop plants able to grow in these adverse conditions assumes importance in this emerging global context. Despite the wealth of information on abiotic stress and stress tolerance in plants, many aspects still remain unclear. One of the effective ways of analyzing a stress response

S. George · B. Usha · A. Parida (✉)
M.S. Swaminathan Research Foundation, Taramani, Chennai 600113, India
e-mail: ajay@mssrf.res.in

Humana Press

is to use model organisms, chosen either for their relative amenability for the study of their tolerance to the stress in question.

Prosopis juliflora, a leguminous phreatophytic perennial deciduous thorny small tree [3], is now the dominant woody plant found in about 45 million hectares of grazing lands worldwide from sea level to 1,500 m. It grows in very hot and dry regions with temperatures as high as 48 °C and annual precipitation of 150–750 mm [4, 5]. The roots penetrate to great depths in the soil, and the tree can grow in a variety of areas including saline, alkaline, sandy, and rocky soils. An introduced species in India, it has subsequently spread aggressively across the country. *P. juliflora* can withstand high leaf-to-air vapor pressure deficit (VPD) by reducing metabolic activity and by effective adjustments in the partitioning of electron flow between assimilation and non-assimilation processes, which in turn imposes a strong limitation on the potential carbon gain [6]. *P. juliflora* is also heavy-metal-tolerant [7], and due to the heavy metal accumulating property, it has been suggested as a green solution for soils contaminated with cadmium, chromium, and copper [8].

Late embryogenesis abundant (LEA) proteins were first identified about 27 years ago in cotton and wheat [9–11]. LEA genes have been cloned from many plant species, and at least six different LEA groups have been defined on the basis of expression pattern and sequence [2, 12, 13]. As their name suggests, they are produced in abundance during late embryo development, comprising up to 4% of cellular protein [14] with maximum expression post-abscission during desiccation [15]. Their expression is linked to the acquisition of desiccation tolerance in orthodox seeds, pollen, and unhydrobiotic plants, but many LEA proteins are induced by cold, osmotic stress of exogenous abscisic acid, or can even be expressed constitutively (e.g., *dhn X* from *Arabidopsis* [16]). A variety of mechanisms have been proposed accounting for the stress-tolerant property of LEA proteins and include protecting cellular structures from the effects of water loss by retention of water, sequestration of ions, and direct protection of other proteins [2, 12, 13, 17].

LEA proteins are generally grouped based on their similarity to prototypical LEA proteins from *Gossypium hirsutum* [18]. In this paper, we report the isolation and characterization of the complementary DNA (cDNA) clone for an atypical LEA protein and its putative promoter sequence from *P. juliflora*. This cDNA clone was obtained from a cDNA library created from drought-stressed leaf tissue of 2-month-old *P. juliflora* plants and happened to be one of the most abundant EST clones in this stress-induced transcripts-enriched cDNA library [19]. The present study was undertaken as a part of the overall program of identifying the stress tolerance mechanisms in *P. juliflora*.

Materials and Methods

Isolation of the cDNA Clone and *In Silico* Analysis

A cDNA library was created from the leaf tissue of drought-stressed 2-month-old *P. juliflora* seedlings [19]. One thousand seven hundred fifty random ESTs were sequenced from this library, and 11 clones showed maximum similarity to a putative atypical LEA protein-like protein from *Ammopiptanthus mongolicus* (Genbank ID: AAW31666). A representative clone (Pj 340, later renamed Pj LEA3—Genbank ID: DQ645423) was completely sequenced. Open reading frames (ORFs) were identified in this cDNA using ORF finder at NCBI (www.ncbi.nlm.nih.gov), and the longest ORF with a predicted start and stop codon was identified as the putative coding region. Protein targeting predictions were done using the Target P 1.1 Server [20, 21]. Molecular mass predictions were made

using Prot Param [22]. HMMTOP [23, 24] was used for making putative transmembrane predictions. Multiple sequence alignments were done using the ClustalW program at NPS (network protein sequence analysis), and phylogenetic analysis was carried out using the program from European Bioinformatics Institute (www.ebi.ac.uk). The conserved domains present in the ORFs were identified using the NCBI conserved domain database [25].

Northern Analysis of Pj LEA3

Total RNA from *P. juliflora* leaf tissue was isolated according to [26]. The expression patterns of *Pj LEA3* in *P. juliflora* leaf tissue under polyethylene glycol (PEG) stress (25%) and H_2O_2 stress (90 mM) at different time points were studied using Northern analysis. Full-length cDNA amplified using UTR specific primers Pj LEA3 5′ UTR F (GAGTTGG GAAGGAATAAGTG) and PjLEA3 3′ UTR R (CCTCTACTTATTTCTAA AACG) was used as the probe (PCR conditions—1 min pre-amplification at 94 °C, 30 PCR cycles of 30 s at 94 °C, 40 s at 55 °C, 1 min and 30 s at 72 °C, a final extension of 10 min at 72 °C). The Northern experiments were carried out as described in [27].

Isolation of the 5′ Upstream Region of Pj LEA3

The 5′ upstream regions of Pj LEA3 were isolated using thermal asymmetric interlaced polymerase chain reaction (TAIL-PCR) [28]. In this, PCR is carried with three long-nested gene-specific primers designed close to the 5′ end of the cDNA [Pj LEA3 GSP1 (CGAGC CATGGTGAGGTGGAAC), Pj LEA3 GSP2 (GATAGGTTAGGGTGATGAATACAAG), Pj LEA3 GSP3 (GAGAGGAACCTTTT CTGATCGG)], and short degenerate primers of arbitrary sequence (primers AD1—NGTCGASWGANAWGAA, AD2—TGWGNAGSANCASAGA, AD3—WGTGNAG WANCANAGA, and AD4—STTGNTASTNCTNTGC). An elaborate thermal cycling program, composed of "super cycles", each consisting of one low stringency cycle and two high stringency cycles, allows only sequence-specific primers to be amplified. The TAIL-PCR reaction products were separated on adjacent lanes in a 1.5% agarose gel, and distinct PCR products showing a difference in size corresponding to the relative positions of the nested gene-specific primers were identified. The specific products obtained were gel-eluted-sequenced. The putative *cis*-acting DNA elements in the isolated 5′ regions were identified using the PLACE database [29].

Promoter Functional Analysis by Transient Reporter Gene Expression

The Pj LEA3 putative promoter was re-amplified from *P. juliflora* genomic DNA using the primers Pj LEA3 PRMF (5′ GTCGAGTGAAAAGAACTTAAAGAG 3′) and Pj LEA3 GSP1 (5′ CGAGCCATGGTGAGGTGGAAC 3′) and cloned in T/A cloning vector (InsTA Clone PCR cloning kit, MBI Fermentas) according to the manufacturer's instructions. After confirming the sequence, Pj LEA3 putative promoter was released from the T/A cloning vector by restriction digestion with *Eco*RI and *Bam*HI and cloned in the promoter fusion vector pCAMBIA 1391z in the *Eco*RI and *Bam*HI sites with *UidA* (GUS) as the reporter gene (pPRM-1391z). pPRM-1391z was transformed into *Agrobacterium* strain LBA4404 by freeze thaw method. Recombinant *Agrobacterium* colonies were selected on YEP medium supplemented with rifampicin (10 mg/ml) and kanamycin (50 mg/ml) followed by colony PCR with Pj LEA3 PRMF and Pj LEA3 GSP1 primers. A single positive *Agrobacterium* colony was resuspended in 100 μl of sterile distilled water and then spread-plated on YEP medium (rifampicin, kanamycin). After 2 days of incubation in dark at 28 °C,

the lawn of culture was scraped off carefully and resuspended in half strength MS medium [30] to obtain an $OD_{595} \cong 1$ to 2. Acetosyringone (100 μM) was added to the suspension and gently mixed for about 10 min. Leaves of greenhouse-grown tobacco cv. petit havana plants were infiltrated in the epidermal region with this culture using a syringe. After 2 days, GUS staining was carried out in infiltrated leaf regions according to [31].

Results

In *silico* Analysis of LEA cDNA

The full-length cDNA pf Pj LEA3 has a total length of 801 bp, and the longest ORF is from 169–450 bp (Genbank ID: DQ645423). Pj LEA3 could be grouped into the class of atypical LEA proteins based on the phylogenetic analysis with 36 LEA proteins from *Gossypium hirsutem* (Fig. 1). The phylogram shows Pj LEA3 clustering with LEA 5 proteins (as per [32]). For a recent comparison of various LEA protein naming systems, see [33].

Atypical LEA proteins share no amino acid sequence similarity with other LEA proteins. In contrast to other LEA proteins, which are very hydrophilic, atypical LEA proteins display a slightly hydropathic character throughout the length of the protein [32]. Unlike typical LEA proteins rich in glycine, Pj LEA3 has alanine as the most abundant amino acid (14%) followed by serine (10.8%) and a predicted molecular weight of 9.9 kDa. Protein

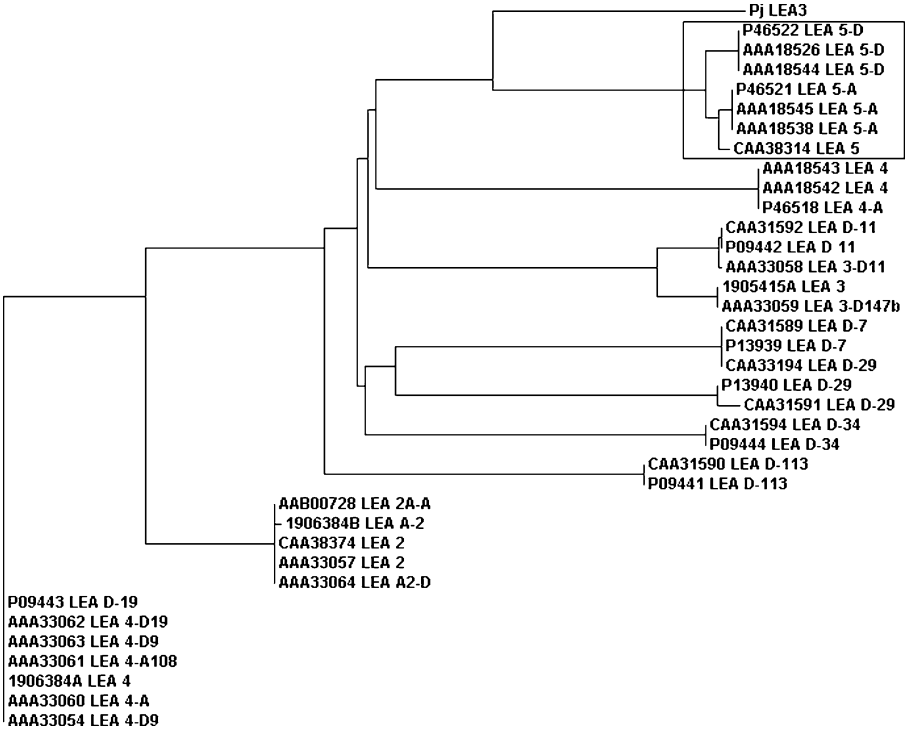

Fig. 1 Phylogenetic analysis of Pj LEA3 with late embryogenesis abundant proteins (*LEA*) from *G. hirsutum*. Genbank accession number followed by annotation is shown in the picture. The atypical LEA proteins from *G. hirsutum* are *boxed*. The analysis shows clustering of Pj LEA3 with atypical LEA proteins

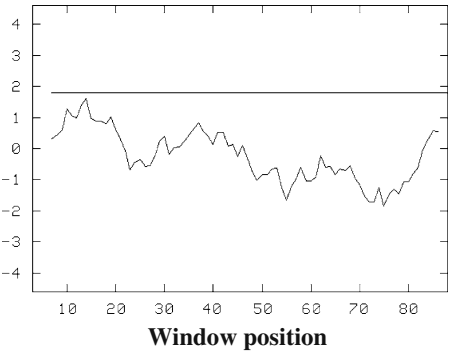

Fig. 2 Hydropathic index analysis of the deduced amino acid sequences of Pj LEA3. Regions above a hydropathy score of zero are hydrophobic. Pj LEA3 showed an average hydropathy score of −0.238. The hydropathy score is according to [34]

targeting predictions indicated that Pj LEA3 might be located in chloroplast/mitochondria. InterProScan results reveal that the protein belongs to the Pfam family LEA_3, PF03242. Kyte and Doolittle [34] hydropathy plot indicated an average negative hydropathy (−0.238), with the N-terminal being more hydrophobic than the C-terminal (Fig. 2). Pj LEA3 was significantly different from other atypical LEA proteins in the database. It showed maximum similarity to putative late-embryogenesis protein-like protein from *Ammopiptanthus mongolicus* (75%) followed by 71% similarity to indole-3 acetic-acid-induced protein ARG2 from *Vigna radiata*. Atypical LEA proteins have been implicated in cellular functions different from those of the hydrophilic LEA proteins [32, 35]. A multiple sequence alignment of Pj LEA3 with first five hits in the BLASTX search revealed a high degree of polymorphism (52.88%) among these proteins. The amino acid segment KEAWVPDPVTGYYRPEN was relatively conserved among all proteins analyzed (Fig. 3).

Northern analysis for Pj LEA3 in *P. juliflora* leaves under 25% PEG stress revealed up-regulation of gene expression at 12 h, a drop in transcript level at 24 h, and an increase

```
                    10         20         30         40         50         60
                     |          |          |          |          |          |
AAW31666      MARSFTNAKVLSVLVLNRFSSSLTRRGYAAAS-AT-QSAKRGGVVSIS--GKMAPKSGEE
AF192758_1    MARSIANAKTFSALVLDGFS----RRGYS-------QSATRGGVASI------APKSGED
P32292        MARSFTNVKVLSALVADGFSNTTTRHGFAAAA-AATQSATRGGASIG---GNMVPKSGEE
Pj LEA3       MARSFSNLNGLSVLLVNAIS----RRGFASSSQGFATTAARVGAASMS--GKVAKKSVEE
ABF29697      MARSFSNAKVISGLIS----EAINGRGFSAVA---SQGAAVSKARSG---AAVMKKTGEE
NP_567231     MARSISNVKIVSAFVSRELSNAIFRRGYAATA---AQGSVSSGGRSGAVASAVMKKKG--
              ****::* : .* ::           :*::          :                 *.

                    70         80         90        100
                     |          |          |          |
AAW31666      KRATTDKVSWVPDPVTGYYKPEN-IKEIDVAELRATLLGKKFN-
AF192758_1    KGVSSYKVSWVPDPVTGYYKPEN-IKEVDVADLRATLLRKKFNN
P32292        KVRGGEKVSWVPDPVTGYYRPEN-TNEIDVADMRATVLGKKFNQ
Pj LEA3       R-VTAEKEAWVPDPVTGYYRPEN-TKEIDVAELRATLLGKK---
ABF29697      VTKTTEKISWVPDPRTGFYRPENVAQEIDAAELRAAFLKKH---
NP_567231     VEESTQKISWVPDPKTGYYRPETGSNEIDAAELRAALLNNKQ--
                    * :***** **:*:**.  :*:*.*::**:.* ::
```

Fig. 3 Multiple sequence alignment of Pj LEA3 with similar proteins from higher plants. *Numbers* indicate position of amino acids from the N terminus. *Asterisk* indicates conserved amino acids across the species: *colon*, strongly similar amino acids; *dot*, weakly similar. The alignment revealed a high degree of polymorphism (52.88%) among the analyzed proteins. Accession nos. AAW31666 (*A. mongolicus*), P32292 (*Oryza sativa*), AF192758_1 (*Glycine max*), ABF29697 (*Populus suaveolens*), NP_567231 (*Arabidopsis thaliana*)

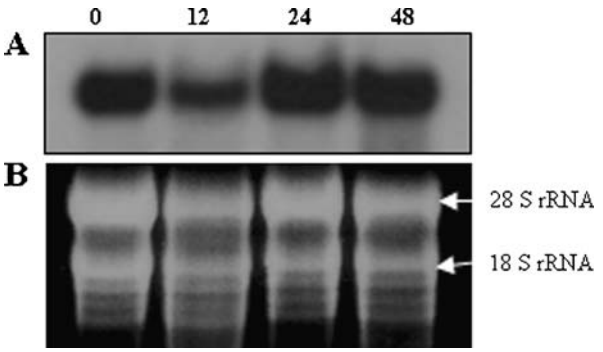

Fig. 4 Northern analysis of Pj LEA3 at 90 mM H_2O_2 stress. Samples were collected at 0, 12, 24, and 48 h of stress application. **a** Transcript analysis of Pj LEA3 in *P. juliflora* leaves under 90 mM H_2O_2 stress. **b** Ethidium-bromide-stained gels prior to transfer

again at 48 h [19]. Under 90 mM H_2O_2 stress, though the transcript level decreased slightly at 12 h after stress, it increased at 24 and 48 h (Fig. 4).

Isolation of the 5′ Upstream Region of Pj LEA3 *In Silico* and Functional Analysis of the Promoter

The TAIL-PCR products of Pj LEA3 were separated on an agarose gel. Three fragments of the sizes 1.4, 1.2, and 0.95 Kb were obtained, and all of them were found to have sequence overlaps with the Pj LEA3 cDNA sequence at their 3′ ends and were hence confirmed to be the specific amplification products. The gel electrophoresis of TAIL-PCR products is shown in Fig. 5. Based on this sequence information, a 1,573-bp putative promoter fragment could be re-amplified from genomic DNA, and the sequence was deposited in Genbank (Genbank ID: EU170656).

Several putative *cis*-acting DNA elements in the 1,573-bp fragment were identified in the upstream sequence and 5′ UTR using the PLACE database [29]. The first ATG of the longest ORF was considered as the putative translation start site (TSS) and numbered +1. The 5′ upstream sequence of Pj LEA3 has a putative TATABOX4 (TATATAA) at position −199. Several ABA-responsive elements, ACGTG, ACGTSSSC, and MACGYGB, could be identified in the sequence [36–38]. *Cis*-acting elements involved in the response to dehydration such as RYCGAC and ACGT were also located in the 5′ upstream sequence of Pj LEA3 [36, 39]. An ethylene-responsive element, AWTTCAAA, was found at −1449 and −832 [40]. GCCGCC, found in many pathogen-responsive genes such as PDF1.2, Thi2.1, and PR4, and CCGAC, a low-temperature responsive element, were also identified [41, 42]. Two MYB recognition elements, WAACCA and YAACKG, and CANNTG, a

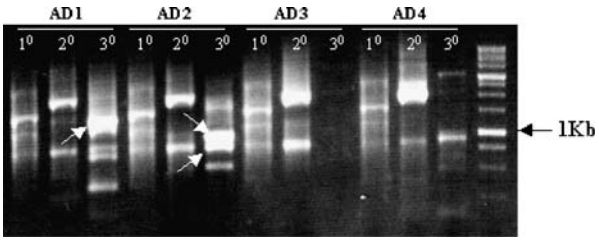

Fig. 5 TAIL-PCR amplification products of Pj LEA3. Primary secondary and tertiary TAIL-PCR reactions using the degenerate primers AD1, AD2, AD, and AD4 and Pj LEA3-specific nested reverse primers are shown. The fragments indicated by *arrows* were cloned, sequenced, and were found to have sequence overlap with the 5′ end of the Pj LEA3 cDNA and thus confirmed to be the specific amplification products

MYC recognition element, were found in the putative promoter [43, 44]. Several *cis*-acting elements involved in tissue-specific expression such as YACT (mesophyll [45]), CTCTT (nodule [46]), AGAAA (pollen [47]), ATATT (root [48]) were also identified in the 5′ upstream sequence of Pj LEA3. The most abundant *cis*-acting element identified in Pj LEA3 was AAAG and is required for binding of Dof proteins [49] and occurred 22 times. GRWAAW, involved in light-regulated transcription [50], and CAANNNNATC, involved in circadian regulation [51], were also identified. ACACNNG, found at −554 and −20, is a core sequence that binds a novel class of bZIP transcription factors, DPBF-1 and 2 (Dc3 promoter-binding factor-1 and 2), and is found in the carrot Dc3 gene promoter. Dc3 expression is normally embryo-specific and also can be induced by ABA [52].

Intense GUS staining was detected in tobacco leaves transiently expressing the Pj LEA3 promoter, *Uid A* construct (pPRM-1391z). This indicated that the isolated Pj LEA3 promoter was a functional promoter in tobacco.

Discussion

LEA proteins were first identified in cotton (*G. hirsutum*) as proteins that accumulated to high levels during the maturation phase of cotton [9]. LEA proteins have been grouped into various families on the basis of sequence similarity. Although significant similarity has not been detected between the members of the different families, the unifying and outstanding feature of most of them is their high hydrophilicity index (greater than 1), a lack or low proportion of Cys and Trp amino acid residues, and a preponderance of Gly (greater than 6%) [53]. The common structural elements among the members of different families indicate that most exist principally as randomly coiled proteins in solution. While structure modeling and structure prediction programs suggest that at least some LEA proteins from particular families contain defined conformations, all hydrophilic LEA proteins studied experimentally have revealed a high degree of unordered structure in solution. This has led them to be considered as intrinsically unstructured proteins [53]. Atypical LEA proteins are those that contain a significantly higher proportion of hydrophobic residues. All LEA proteins with a higher content of hydrophobic residues than typical LEA proteins are included in this group; thus, this group incorporates non-homologous proteins and include pcC27-45 (*Craterostigma*; P22241), cotton LEA 14-A (NP_171654), soybean D95 (P46519), and tomato Lemmi9 and ER5 (CAA93156 and AAB96796, respectively) [32, 35]. Given their physicochemical properties, atypical LEA proteins are not soluble after boiling, suggesting that they adopt a globular conformation [32, 54, 55]. Although little is known about this set of proteins, the available data indicate that their transcripts accumulate during the late stage of seed development and in response to stress conditions, such as drought, UV light, salinity, cold, and wounding [53].

Pj LEA3 was classified as an atypical LEA protein based on phylogenetic analysis with LEA proteins from *G. hirsutum* and the amino acid composition different from typical LEA proteins. Protein targeting predictions indicated that Pj LEA3 might be located in chloroplast/mitochondria, like four atypical LEA proteins reported from *Arabidopsis* [33]. Atypical LEA proteins have been implicated in cellular functions different from those of the hydrophilic LEA proteins [32, 35]. ClustalW of Pj LEA3 with the first few BLAST hits reveals the highly unconserved nature of this protein. Northern analysis for Pj LEA3 in *P. juliflora* leaves under PEG stress revealed up-regulation at 12 h, a drop in transcript level at 24 h, and an increase again at 48 h [19]. Kim et al. [56] reported the up-regulation of an atypical LEA protein from *Capsicum annum* (*CaLEA6*) following treatment with PEG, ABA, and NaCl. Overexpression of this gene in tobacco plants resulted in enhanced

tolerance to dehydration and NaCl. Thus, the up-regulation of Pj LEA3 could be of importance under water stress. Under H_2O_2 stress, though the transcript level decreased slightly at 12 h after stress, it increased at 24 and 48 h. The up-regulation of Pj LEA3 under H_2O_2-induced oxidative stress indicates that Pj LEA3 probably could impart stress tolerance through the removal of reactive oxygen species.

The precise function of LEA proteins is unknown, but they are assumed to protect cellular or molecular structures from the damaging effects of water loss; a number of putative mechanisms have been proposed, including hydration buffering, ion sequestration, direct protection of other proteins or membranes, renaturation of unfolded proteins, or prevention of protein aggregation due to water stress [12, 13, 57]. Interestingly, only two other late embryogenesis abundant proteins were found in the *P. juliflora* cDNA library, both singletons, one of them being a dehydrin.

Several putative *cis*-acting DNA elements were identified in the upstream sequence and 5′ UTR of the Pj LEA3, including CAATBOX, GATA BOX, and TATABOX. Several ABA-responsive elements and *cis*-acting elements involved in response to dehydration were identified at various positions in the sequence. The ACGT element is involved in early response to drought of *erd1*, an *Arabidopsis* gene [36]. These elements might contribute to the up-regulation of Pj LEA3 at 12 h of PEG stress in *P. juliflora* leaves. An ethylene-responsive element, a *cis*-element found in many pathogen-responsive genes, a low temperature responsive element, several Myb and Myc recognition elements, etc. were also identified. Several *cis*-acting elements involved in tissue-specific expression were found in the 5′ upstream sequence of Pj LEA3. The most abundant *cis*-acting element was AAAG, required for binding of Dof proteins. GRWAAW involved in light-regulated transcription and CAANNNNATC involved in circadian regulation were also identified. The Pj LEA3 promoter was found to be functional in tobacco as revealed by transient reporter gene expression. Further characterization of this promoter would throw light into the possible functions of this gene. The results of our studies could be used as a starting point for generation of crop plants tolerant to abiotic stresses.

Acknowledgment This work was carried out with a grant from Department of Biotechnology (DBT), Government of India. Ms. Suja George is a Senior Research Fellow of the Council of Scientific and Industrial Research (CSIR), India.

References

1. Boyer, J. S. (1982). *Science*, *218*, 443–448. doi:10.1126/science.218.4571.443.
2. Bray, E. A., Bailey-Serres, J., & Weretilnyk, E. (2000). In W. Gruissem, B. Buchannan, & R. Jones (Eds.), *Responses to abiotic stresses* pp. 1158–1249. Rockville, MD: American Society of Plant Physiologists.
3. Burkart, A., & Simpson, B. B. (1977). In B. B. Simpson (Eds.), *Mesquite: Its biology in two desert ecosystems* pp. 201–215. Stroudsburg, Pennsylvania, USA: Dowden, Hutchinson and Ross.
4. Drake, H. (1993). *Trees for dry lands* p. 370. New York: International Scientific Publishing.
5. Geilfus, F. (1994). In *Enda-Caribe-Centro Agronómico Tropical de Investigación y Enseñanza* (vol. 2, p. 597).
6. Shirke, P. A., & Pathre, U. V. (2004). *Journal of Experimental Botany*, *55*(405), 2111–2120. doi:10.1093/jxb/erh229.
7. Sinha, S., Rai, U. N., Bhatt, K., Pandey, K., & Gupta, A. K. (2005). *Environmental Monitoring and Assessment*, *102*, 447–457. doi:10.1007/s10661-005-6397-4.
8. SenthilKumar, P., Prince, W. S., Sivakumar, S., & Subbhuraam, C. V. (2005). *Chemosphere*, *60*(10), 1493–1496. doi:10.1016/j.chemosphere.2005.02.022.
9. Dure, L., Greenway, S. C., & Galau, G. A. (1981). *Biochemistry*, *20*, 4162–4168. doi:10.1021/bi00517a033.

10. Grzelczak, Z. F., Sattolo, M. H., Hanley-Bowdoin, L. K., Kennedy, T. D., & Lane, B. G. (1982). *Canadian Journal of Biochemistry*, *60*, 389–397.
11. Galau, G. A., Hughes, D. W., & Dure, L. (1986). *Plant Molecular Biology*, *7*, 155–170.
12. Bray, E. A. (1993). *Plant Physiology*, *103*, 1035.
13. Cumming, A. C. (1999). In P. R. Shewry, & R. Casey (Eds.), *Seed proteins* (pp. 753–780). Dordrecht: Kluwer
14. Roberts, J. K., DeSimone, N. A., Lingle, W. L., & Dure, L. (1993). *The Plant Cell*, *5*, 769–780.
15. Hughes, D. W., & Galau, G. A. (1989). *Genes & Development*, *3*, 358–369. doi:10.1101/gad.3.3.358.
16. Welin, V., Olson, A., Nylander, M., & Palva, E. T. (1994). *Plant Molecular Biology*, *26*, 131–144. doi:10.1007/BF00039526.
17. Ingram, J., & Bartels, D. (1996). *Annual Review of Plant Biology*, *47*, 377–403. doi:10.1146/annurev.arplant.47.1.377.
18. Wise, M. J. (2003). *BMC Bioinformatics*, *4*, 52 1040.
19. George, S., Venkataraman, G., & Parida, A. (2007). *Genome*, *50*(5), 470–478. doi:10.1139/G07-014.
20. Emanuelsson, O., Nielsen, H., Brunak, S., & von Heijne, G. (2000). *Journal of Molecular Biology*, *300*, 1005–1016. doi:10.1006/jmbi.2000.3903.
21. Nielsen, H., Engelbrecht, J., Brunak, S., & von Heijne, G. (1997). *Protein Engineering*, *10*, 1–6. doi:10.1093/protein/10.1.1.
22. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., et al. (2005); In J. M. Walker (Ed.), *The proteomics protocols handbook* (pp. 571–607). New York: Humana.
23. Tusnády, G. E., & Simon, I. (1998). *Journal of Molecular Biology*, *283*, 489–506. doi:10.1006/jmbi.1998.2107.
24. Tusnády, G. E., & Simon, I. (2001). *Bioinformatics (Oxford, England)*, *17*, 849–850. doi:10.1093/bioinformatics/17.9.849.
25. Bauer, M. A., Anderson, J. B., Derbyshire, M. K., DeWeese-Scott, C., Gonzales, N. R., & Gwadz, M. (2007). *Nucleic Acids Research*, *35*, 237–240. doi:10.1093/nar/gkl951.
26. Chomczynski, P., & Sacchi, N. (1987). *Analytical Biochemistry*, *162*, 156–159. doi:10.1016/0003-2697(87)90021-2.
27. Mehta, P. A., Sivaprakash, K., Parani, M., Venkataraman, G., & Parida, A. K. (2005). *Theoretical and Applied Genetics*, *110*, 416–424. doi:10.1007/s00122-004-1801-y.
28. Liu, Y. G., Mitsukawa, N., Oosumi, T., & Whittier, R. F. (1995). *The Plant Journal*, *8*, 457–463. doi:10.1046/j.1365-313X.1995.08030457.x.
29. Higo, K., Ugawa, Y., Iwamoto, M., & Korenaga, T. (1999). *Nucleic Acids Research*, *27*, 97–300. doi:10.1093/nar/27.1.297.
30. Murashige, T., & Skoog, F. (1962). *Physiologia Plantarum*, *15*, 473–497. doi:10.1111/j.1399-3054.1962.tb08052.x.
31. Jefferson, R. A. (1987). *Plant Molecular Biology Reporter*, *5*, 387–405. doi:10.1007/BF02667740.
32. Galau, G. A., Wang, H. Y.-C., & Hughes, D. W. (1993). *Plant Physiology*, *101*, 695–696. doi:10.1104/pp.101.2.695.
33. Hundertmark, M., & Hincha, D. K. (2008). *BMC Genomics*, *9*, 118. doi:10.1186/1471-2164-9-118.
34. Kyte, J., & Doolittle, R. (1982). *Journal of Molecular Biology*, *157*, 105–132. doi:10.1016/0022-2836(82)90515-0.
35. Zegzouti, H., Jones, B., Marty, C., Lelievre, J. M., Latche, A., & Pech, J. C. (1997). *Plant Molecular Biology*, *35*(6), 847–854. doi:10.1023/A:1005860302313.
36. Simpson, S. D., Nakashima, K., Narusaka, Y., Seki, M., Shinozaki, K., & Yamaguchi-Shinozaki, K. (2003). *Plant Journal*, *33*, 259–270. doi:10.1046/j.1365-313X.2003.01624.x.
37. Marcotte, W. R., Russell, S. H., & Quatrano, R. S. (1989). *The Plant Cell*, *1*, 969–976.
38. Kaplan, B., Davydov, O., Knight, H., Galon, Y., Knight, M. R., & Fluhr, R. (2006). *The Plant Cell*, *18*, 2733–2748. doi:10.1105/tpc.106.042713.
39. Xue, G. P. (2002). *Nucleic Acids Research*, *30*, e77. doi:10.1093/nar/gnf076.
40. Itzhaki, H., Maxson, J. M., & Woodson, W. R. (1994). *Proceedings of the National Academy of Sciences of the United States of America*, *91*, 8925–8929. doi:10.1073/pnas.91.19.8925.
41. Brown, R. L., Kazan, K., McGrath, K. C., Maclean, D. J., & Manners, J. M. (2003). *Plant Physiology*, *132*, 1020–1032. doi:10.1104/pp.102.017814.
42. Baker, S. S., Wilhelm, K. S., & Thomashow, M. F. (1994). *Plant Molecular Biology*, *24*, 701–713. doi:10.1007/BF00029852.
43. Abe, H., Urao, T., Ito, T., Seki, M., Shinozaki, K., & Yamaguchi-Shinozaki, K. (2003). *The Plant Cell*, *15*, 63–78. doi:10.1105/tpc.006130.
44. Hartmann, U., Sagasser, M., Mehrtens, F., Stracke, R., & Weisshaar, B. (2005). *Plant Molecular Biology*, *57*, 155–171. doi:10.1007/s11103-004-6910-0.

45. Gowik, U., Burscheidt, J., Akyildiz, M., Schlue, U., Koczor, M., & Streubel, M. (2004). *The Plant Cell*, *16*, 1077–1090. doi:10.1105/tpc.019729.
46. Stougaard, J., Jorgensen, J. E., Christensen, T., Kuhle, A., & Marcker, K. A. (1990). *Molecular & General Genetics*, *220*, 353–360. doi:10.1007/BF00391738.
47. Bate, N., & Twell, D. (1998). *Plant Molecular Biology*, *37*, 859–869. doi:10.1023/A:1006095023050.
48. Elmayan, T., & Tepfer, M. (1995). *Transgenic Research*, *4*, 388–396. doi:10.1007/BF01973757.
49. Yanagisawa, S. (2000). *The Plant Journal*, *21*, 281–288. doi:10.1046/j.1365-313x.2000.00685.x.
50. Zhou, D. X. (1999). *Trends in Plant Science*, *4*, 210–214. doi:10.1016/S1360-1385(99)01418-1.
51. Piechulla, B., Merforth, N., & Rudolph, B. (1998). *Plant Molecular Biology*, *38*, 655–662. doi:10.1023/A:1006094015513.
52. Kim, S. Y., Chung, H. J., & Thomas, T. L. (1997). *The Plant Journal*, *11*, 1237–1251. doi:10.1046/j.1365-313X.1997.11061237.x.
53. Battaglia, M., Olvera-Carrillo, Y., Garciarrubio, A., Campos, F., & Covarrubias, A. A. (2008). *Plant Physiology*, *148*, 6–24. doi:10.1104/pp.108.120725.
54. Baker, J., Steele, C., & Dure, L. (1988). *Plant Molecular Biology*, *11*, 277–291. doi:10.1007/BF00027385.
55. Cuming, A. C. (1999). LEA proteins. In R. Casey & P. R. Shewry (Eds.), *Seed proteins* (pp. 753–780). Dordrecht, The Netherlands: Kluwer.
56. Kim, H. S., Lee, J. H., Kim, J. J., Kim, C. H., Jun, S. S., & Hong, Y. N. (2005). *Gene*, *3*(344), 115–123. doi:10.1016/j.gene.2004.09.012.
57. Goyal, K., Walton, L. J., & Tunnacliffe (2005). *Biochemical Journal*, *388*, 151–157, (15).