Received: 15 June 2021 | Accepted: 16 August 2021

DOI: 10.1111/ppl.13544

SPECIAL ISSUE ARTICLE

Physiologia Plantarum

Functional characterization of late embryogenesis abundant genes and promoters in pearl millet (*Pennisetum glaucum* L.) for abiotic stress tolerance

Kummari Divya[1] | **Sudhakar Reddy Palakolanu**[1] | **Polavarapu Kavi Kishor**[2] | **Aishwarya Shankhapal Rajesh**[1] | **Vincent Vadez**[1] | **Kiran K. Sharma**[1] | **Pooja Bhatnagar Mathur**[1]

[1]International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad, Andhra Pradesh, India

[2]Department of Biotechnology, Vignan's Foundation for Science, Technology & Research Deemed to be University, Vadlamudi, Guntur, Andhra Pradesh, India

Correspondence
Palakolanu Sudhakar Reddy and Pooja Bhatnagar Mathur, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, Hyderabad, 502324, India.
Email: p.sudhakarreddy@cgiar.org and p.bhatnagar@cgiar.org

Funding information
Department of Science and Technology, Ministry of Science and Technology, India, Grant/Award Number: IFALSPA-06; Core Grant Research, Grant/Award Number: EMR/2016/006726

Edited by: S. Penna

Abstract

Late embryogenesis abundant (*LEA*) genes display distinct functions in response to abiotic stresses in plants. In pearl millet (*Pennisetum glaucum* L.), a total of 21 *PgLEA* genes were identified and classified into six groups including *LEA1*, *LEA2*, *LEA3*, *LEA5*, *LEA7*, and dehydrins (*DHN*). Open reading frames (ORFs) of *PgLEA*s range from 291 bp (*PgLEA1-1*) to 945 bp (*PgLEA2-11*) and distributed randomly among the seven chromosomes. Phylogenetic analysis revealed that all PgLEA proteins are closely related to sorghum LEA proteins. The *PgLEAs* were found to be expressed differentially under high progressive vapor pressure deficit (VPD), *PgLEA7* was significantly expressed under high VPD and was selected for functional validation. In silico analysis of the *PgLEA* promoter regions revealed abiotic stress-specific *cis*-acting elements such as ABRE, CCAAT, MYBS, and LTRE. Based on the type of motifs, *PgLEAPC* promoter (758 bp), its deletion 1 (PgLpd1, 349 bp) and deletion 2 (PgLpd2, 125 bp) were cloned into the plant expression vector pMDC164 having the promoter-less *uid*A gene. All the three plant expression vectors were introduced into tobacco through *Agrobacterium tumefaciens*-mediated transformation to obtain T_1 and T_2 generations of transgenic plants. Based on expression of the *uid*A gene, tissue-specific expression was observed in mature stems, roots and seedlings of *PgLEAPC* and *PgLpd1* carrying transgenics only. While the transgenic *PgLEAPC* plants displayed significantly higher *uid*A expression in the stem and root tissues under salt, drought, heat, and cold stresses, very low or no expression was observed in *PgLpd1* and *PgLpd2* transgenics under the tested stress conditions. The results of this study indicate that the complete promoter of *PgLEAPC* plays a role in developing abiotic stress tolerance in plants.

1 | INTRODUCTION

Higher plants display a plethora of physiological, biochemical, and molecular changes in response to harsh environmental conditions (Soltis & Soltis 2003). As a response to water deficit, plants accumulate higher levels of soluble sugars and other compatible solutes like proline, trehalose, glycine betaine and mannitol These molecules help in osmotic adjustment, to stabilize proteins and enzymes, and mitigate the damages caused by the reactive oxygen species (ROS; Liu et al. 2019; Reddy et al. 2015a, 2015b). Several

biotechnological approaches have been implicated in developing resistance to abiotic stress in crop plants (Bhatnagar-Mathur et al. 2008). The evolution of LEA proteins is one of the adaptive mechanisms acquired by plants for protection against harsh environmental conditions (Kishor et al. 2020). LEA proteins mainly consist of hydrophilic amino acids organized in a repeated sequence (e.g., Gly and Lys), conferring hyper-hydrophilicity and thermal stability that are highly expressed during the late stages of embryo development (Wang et al. 2003), mainly in cotyledons, panicles (Zhang et al. 2002), stems, leaves, and roots (Federspiel 2000).

Genome-wide characterization of *LEA* genes in different crops has been reported. Depending on the similarities in the sequence as well as conserved motif regions, 34 *LEA* genes have been identified in *Oryza sativa* (Wang et al. 2007), 51 in *Arabidopsis* (Battaglia et al. 2008), 53 in *Populus* (Lan et al. 2013), 30 in *Prunus mume* (Du et al. 2013), 36 in *Brachypodium distachyon* (Filiz et al. 2013), 29 in *Solanum tuberosum* (Charfeddine et al. 2015), 27 in tomato (Cao & Li 2015), 108 in *Brassica napus* (Liang et al. 2016), and 68 in *Sorghum bicolor* (Nagaraju et al. 2019). While the overexpression of *LEA* genes has been demonstrated in both model and crop plants under different abiotic stress conditions, the *OsLEA3-1* gene in rice showed a significant increase under drought stress without any yield penalty (Xiao et al. 2007). Sweet potato *LEA14* (*IbLEA*14) displayed osmotic- and salt stress tolerance in transgenic calluses of sweet potato (Park et al. 2011). Transgenic *Arabidopsis* overexpressing *RcLEA* exhibited better growth after exposure to high and low-temperature regimen and low levels of peroxide (Zhang et al. 2014). The modification of *SiLEA14* expression demonstrated improved abiotic stress tolerance in crops (Wang et al. 2014). While these studies infer that *LEA* genes are implicated in abiotic stress tolerance, it must be noted that several studies employing transgenics for mechanistic studies have often used constitutive promoters for gene expression studies where yield penalty has been noticed when compared to stress-inducible promoters (Bhatnagar-Mathur et al. 2007; Divya et al. 2020; Du et al. 2016). While the constitutive overexpression of genes may impede the energy necessary for normal growth and RNA biosynthesis (Malnoy et al. 2006), stress-inducible promoters limit gene expressions on a need-basis, thereby reducing the adverse effects on crop yield. This necessitates studies on the identification, validation and use of stress-inducible promoters in efforts towards unraveling the molecular mechanisms associated with signaling pathways under abiotic stress (Bhatnagar-Mathur et al. 2008; Mehrotra et al. 2011).

Responses of transpiration to increasing VPD have been linked to final yield of the crop under terminal drought stress (Vadez et al. 2014), where limitations in transpiration demand under high VPD conditions are genotype-specific in pearl millet and also linked to water-deficit adaptation mechanisms (Kholova & Vadez 2013). Therefore, to understand the expression profile of the *PgLEA* family genes, genotypes that exhibit contrasting VPD were selected. Further, stress-responsive promoters have been deployed previously for eliminating the negative effects of constitutive promoters under abiotic stress conditions (Banerjee et al. 2013; Bhatnagar-Mathur et al. 2007; Yan et al. 2015). *LEA* promoters like *PpLEA-1* pro from *Physcomitrella* are highly active in response to osmotic stress (Kamisugi & Cuming 2005), while ABA- and stress-inducible group 4 *LEA* family of promoters from *B. napus* (*Bnlea4-1*) displayed increased salt and drought tolerance (Dalal et al. 2009). The *CuLEA* promoter from *Citrus unshiu* has been shown to confer fruit-preferential and stress-inducible gene expression in *Arabidopsis* (Kim et al. 2011). The *LEA* genes have not yet been identified from pearl millet and reports regarding their systematic characterization are scant (Reddy et al. 2012), besides a few reports on the functional analysis of the promoters till date (Jin et al. 2019). In the present study, members of the *LEA* gene family from pearl millet have been identified and characterized following functional analysis of the *PgLEAPC* promoter region in transgenic tobacco plants, including two of its deletion constructs individually using the *uid*A reporter gene. Three independent homozygous transgenic lines of tobacco in T_2 generation where the *uid*A gene was regulated by *PgLEA* promoters were functionally analyzed in different tissues under diverse abiotic stresses and compared with their wild type controls.

2 | MATERIALS AND METHODS

2.1 | Identification of *LEA* genes in pearl millet

LEA genes in pearl millet were identified based on homology with LEA proteins from sorghum and rice using the NCBI BLAST search against the pearl millet genome (http://cegsb.icrisat.org/ipmgsc/) (Varshney et al. 2017). Coding sequences (CDS) and chromosomal localization of *PgLEA* genes were retrieved from the pearl millet genome. Intron and exon structures of PgLEA family genes were obtained by aligning the CDS with their respective genomic sequences using Gene structure display server (http://gsds.cbi.pku.edu.cn/). The number of amino acids, molecular weight (MW) and isoelectric point (pI) were obtained using ExPASy (http://www.expasy.org/tools/) and Mac Vector assembly program (V17.1). *PgLEA* genes were analyzed for the domains using the Pfam (http://pfam.sanger.ac.uk/search) and Simple Modular Architecture Research Tool (SMART) (http://smart.embl-heidelberg.de/) databases to confirm that all genes retrieved belong to the *LEA* gene family. ProtParam tool (http://web.expasy.org/protparam/) was used to calculate the GRAVY (grand average of hydropathy) values. Subcellular location was predicted by the Wolf PSort programme (Horton et al. 2007). The physiochemical properties were calculated using ExPASy. Conserved motifs present in PgLEA protein subfamilies were predicted using MEME (Multiple Expectation Maximization for Motif Elicitation; http://alternate.meme-suite.org/) software (Bailey et al. 2009). An unrooted phylogenetic tree of the full-length *LEA* protein sequences was constructed using MEGA 6 with the Neighbor Joining (NJ) method and bootstrap analysis conducted using 1000 replicates. The protein as well as CDS of respective orthologs and paralogs were used to calculate synonymous and non-synonymous substitution rates using PAL2NAL.

2.2 | *PgLEAPC* promoter cloning and *cis*-motif analysis

Sequence information of the *PgLEA* promoter (Reddy et al. 2012) was used for cloning the *PgLEAPC* promoter region from the pearl millet genomic DNA (Gene bank accession No. JN380919), which belongs to the LEA 7 group. Based on the type of *cis*-motifs, the *PgLEA* gene promoters including *PgLEAPC*, *PgLpd1*, and *PgLpd2* measuring 758, 349 and 125 bp, respectively, were cloned from the genomic DNA, sequenced, and subjected to in silico sequence analysis using the PlantCARE (Lescot et al. 2002) and PLACE (Higo et al. 1999) databases. Based on the type of *cis*-motifs and location, PCR primers were designed for the *PgLEAPC* full-length promoter and the successive 5′ truncations. Promoter regions of the *PgLEAPC*, *PgLEALpd1*, and *PgLEALpd2* were PCR amplified and cloned into pCRGW8-TOPOTA vector (Invitrogen) and sequenced. *PgLEAPC*, *PgLpd1*, and *PgLpd2* promoters were cloned into pMDC164 plant expression vector individually using Gateway cloning upstream of the *uid*A coding region. The cloned constructs *PgLEA-uid*A-Nos, *PgLpd1-uid*A-Nos, and *PgLpd2-uid*A-Nos in pMDC164 were transferred into *Agrobacterium tumefaciens* (strain EHA105) by electroporation and recombinant colonies containing *PgLEAPC-uid*A-Nos, *PgLpd1-uid*A-Nos, or *PgLpd2*-uidA-Nos were used for the genetic transformation of tobacco.

2.3 | Development of transgenic tobacco plants

An *A. tumefaciens* vector containing *PgLEAPC-uid*A-Nos, *PgLpd1-uid*A-Nos or *PgLpd2-uid*A-Nos were used individually for the genetic transformation of tobacco (*Nicotiana tabacum* variety Xanthi) with the respective plant transformation vector (Divya et al., 2019, 2020). A total of 10, 8, and 34 hygromycin-resistant T_0 transgenic tobacco plants were generated with *PgLEAPC*, *PgLpd1*, and *PgLpd2* promoters, respectively. Genomic DNA was isolated from the leaves of 20-day-old transgenic tobacco plants using the Nucleo Spin plant II midi kit (Macherey-Nagel) following the manufacturer's protocol. Integration of the transgene was confirmed with junction primers of promoter (T-LEAPC-F/T-LeaPD1-F/T-LeaPD2-F) and GUS gene (*uid*A-R) and *hygromycin phosphotransferase* gene specific primers (Hpt-F and Hpt-R). Confirmed transgenic plants were grown to maturity and their seeds were resown to obtain the T_1 and T_2 generation plants homozygous for the transgene through PCR. DNA isolations were carried out from 30-day-old T_1 plants of each construct and PCR amplification carried out with the junction primers and hpt primers. From the identified transgenic plants, 10 (*PgLEAPC)*, eight (*PgLpd1*), and 34 (*PgLpd2*) plants from events number 6, 7, and 9, respectively, were found to be PCR positive. To obtain the T_2 generation plants, 100 transgenic tobacco seeds harboring the three promoters-*uidA* constructs in the T_1 generation were surface sterilized individually and plated on MS basal medium (Murashige & Skoog 1962) containing 3% sucrose and 50 mg/L hygromycin. To confirm the Mendelian inheritance ratio, transgenic seeds were germinated on MS medium. Three events in the T_2 generation from each construct, *PgLEA*: 2–5, 2–9, 2–10; *PgLpd1*: 1–1, 6–11, 6–7; *PgLpd2*: 2–6, 27–6, 27–11 were used for validation of promoter regions by histochemical and qRT-PCR analysis.

2.4 | Abiotic stress treatments imposed on transgenic tobacco plants

Thirty-day-old T_2 generation transgenic tobacco plants with *LEAPC*, *Lpd1*, and *Lpd2* promoters were subjected to extreme temperatures, salt, and drought stresses. For heat and cold, plants were incubated for 4 h at 45 and 4°C, respectively. For salt stress, the selected plants were irrigated with 250 mM NaCl solution for 48 h. To impose water deficit, water was held back for 5 days in the selected transgenic plants, while the unstressed control plants were maintained under well-watered conditions. Leaf, stem, and root tissues were collected from the respective control and stress-treated plants for histochemical assays and qRT-PCR analysis.

2.5 | Histochemical GUS staining of transgenic tobacco plants

Histochemical staining for β-glucuronidase (*uid*A) activity was performed in the wild type and stress-treated transgenic plants as described earlier (Jefferson et al. 1987). All the tissue samples were incubated with GUS staining solution (100 mM sodium phosphate [Na_2HPO_4] buffer pH 7.0, 0.5 mM potassium ferrocyanide, 0.5 mM potassium ferricyanide, 0.1% Triton X-100 and 1 mg/L X-Gluc [5-bromo-4-chloro-3-indolyl β-D-glucuronide]) for 20 min and incubated overnight at 37°C in dark. The stained tissues were cleared in ethanol: acetic acid (3:1; v/v) to remove the pigments and chlorophyll before histological analysis and documentation.

2.6 | Expression analysis using quantitative real-time PCR

To understand the expression profile of the *PgLEA* family of genes, genotypes ICMR1122 and ICMR1152 that exhibit contrasting VPD, as described previously (Reddy et al. 2017), were selected for the present study. Quantitative real-time PCR (qRT-PCR) was performed on leaf and root tissues of these genotypes following exposure to VPD conditions. Total RNA was isolated from the stress-exposed and control transgenic tobacco samples using RNA extraction kit (Qiagen) according to the manufacturer's instructions. Purity and integrity of RNA samples were analyzed by agarose gel electrophoresis and by taking the ratio at A260/A230 and A260/A280. Primers for real-time PCR were designed using Primer 3 software (Table S1) and reactions performed in a total volume of 10 μl containing 1 μl of RNA (30 ng), 400 nM of each primer, 5 μl of 2X one-step SYBR RT-PCR buffer 4 (Takara) and 0.5 μl of Prime Script One Step Enzyme Mix 2 (Takara) and made to 10 μl with RNase-free H_2O. The qRT-PCR reactions of all samples were performed by following standard thermal profile:

TABLE 1 In silico characterization of *PgLEA* gene family

S. no.	Gene name	Acc. No.	Chromosome localization	Nucleotide		Protein						Cellular localization
				ORF (bp)	Exon	Protein (aa)	M.Wt	pI	Instability index	GRAVY	Aliphatic index	
1	PgLEA1-1	Pgl_GLEAN_10012637	3	291	1	96	10090.44	9.33	13.69	−1.089	31.98	Mito
2	PgLEA2-1	Pgl_GLEAN_10003225	1	669	2	222	23438.85	7.51	40.6	0.153	89.23	Cyto
3	PgLEA2-2	Pgl_GLEAN_10003471	1	660	1	219	23957.67	9.11	42.41	0.135	94.7	Cyto
4	PgLEA2-3	Pgl_GLEAN_10018806	1	522	2	173	19017.76	5.07	35.49	−0.21	85.66	Nucl
5	PgLEA2-4	Pgl_GLEAN_10038309	1	642	1	213	23176.49	8.28	44.46	0.139	98.36	Chlo
6	PgLEA2-5	Pgl_GLEAN_10038314	1	624	1	207	22224.45	9.3	42.39	0.01	87.54	Chlo
7	PgLEA2-6	Pgl_GLEAN_10025348	3	585	1	194	21188.42	9.58	29.17	0.209	94.02	Mito
8	PgLEA2-7	Pgl_GLEAN_10026514	3	666	2	221	23876.51	7.65	42.55	0.263	100.05	Extr
9	PgLEA2-8	Pgl_GLEAN_10021053	6	525	2	174	19085.79	5.13	37.62	−0.11	91.32	Cyto
10	PgLEA2-9	Pgl_GLEAN_10028629	6	279	2	92	9704.16	9.86	30.5	−0.329	66.09	Chlo
11	PgLEA2-10	Pgl_GLEAN_10001929	7	612	1	203	22765.38	9.1	42.41	−0.054	86.95	Chlo
12	PgLEA2-11	Pgl_GLEAN_10022312	7	945	1	314	35109.89	4.85	29.88	−0.415	91.82	Cyto
13	PgLEA3-1	Pgl_GLEAN_10005440	1	294	2	97	10345.87	6.23	37.42	−0.066	82.58	Chlo
14	PgLEA3-2	Pgl_GLEAN_10020225	4	630	1	209	23184.67	9.14	39.35	0.072	96.51	Chlo
15	PgLEA5-1	Pgl_GLEAN_10018265	1	282	2	93	9922.66	5.45	38.47	−1.311	40.97	Nucl
16	PgLEA5-2	Pgl_GLEAN_10018263	1	462	2	153	16416.61	5.91	51.64	−1.422	33.2	Nucl
17	PgLEA7	JN380919	-	525	-	175	19123.83	6.37	40.21	−0.327	76.91	Cyto
18	PgDHN-1	Pgl_GLEAN_10002256	2	309	1	102	11655.7	6.7	40.76	−2.158	3059	Nucl
19	PgDHN-2	Pgl_GLEAN_10025327	3	897	2	298	32502.03	5.75	50.39	−1.238	55.74	Nucl
20	PgDHN-3	Pgl_GLEAN_10023980	5	402	2	133	13865.1	7.07	33.96	−1.234	42.63	Nucl
21	PgDHN-4	Pgl_GLEAN_10021551	6	522	2	173	17355.84	9.07	15.26	−1.085	27.05	Nucl

42°C for 5 min and 95°C for 10 s (reverse transcription) followed by 40 cycles of 15 s at 95°C, 15 s at 62°C with fluorescent signal recording and 15 s at 72°C. After 40 cycles, amplicons dissociation curves were measured by heating at 95°C and fluorescence measured within 20 min. The 2 (-Delta Delta C[T]) method (Livak & Schmittgen 2001) was used to analyze the expression levels of *uid*A by normalizing with corresponding control tobacco samples and with *UBC* and *EF-1α* reference genes (Schmidt & Delaney 2010). In pearl millet the relative expression levels of *PgLEA* genes in response to low and high VPD conditions was estimated using qbase+ software (Hellemans et al. 2007) by normalizing with *PgMDH* and *PgUBC* reference genes (Reddy et al. 2015a, 2015b).

2.7 | Statistical analysis

Results were expressed as the means ± SE and the data represent average values from three independent experiments. T_1 and T_2 generation seeds obtained from *PgLEAPC*, *PgLpd1*, and *PgLpd2* transgenic tobacco plants were collected after selfing and germinated on the MS medium containing 50 mg/L hygromycin for the selection of transgenic events. Total number of seeds kept for germination, seedlings germinated and not germinated were counted and tested for the Mendelian inheritance pattern (χ^2; Pearson 1900). Normalized expression of the *uid*A gene in different tobacco tissues was normalized against *NtUBC* and *NtEF-1α* reference genes using 2 (-Delta Delta C[T]) method. The relative expression of *uid*A gene under different abiotic stress conditions was analyzed using 2 (-Delta Delta C[T]) method, where the data was normalized against two reference genes and control in leaf, stem and root tobacco samples collected at the same time. Expression analysis of all the qRT-PCR experiments were analyzed with CoStat version 6.204 (Cohort Software) and One-way ANOVA carried out to test between the transgenic events of three promoters. The means were compared using the Tukey–Kramer test as well as LSD (at $P < 0.05\%$ significance).

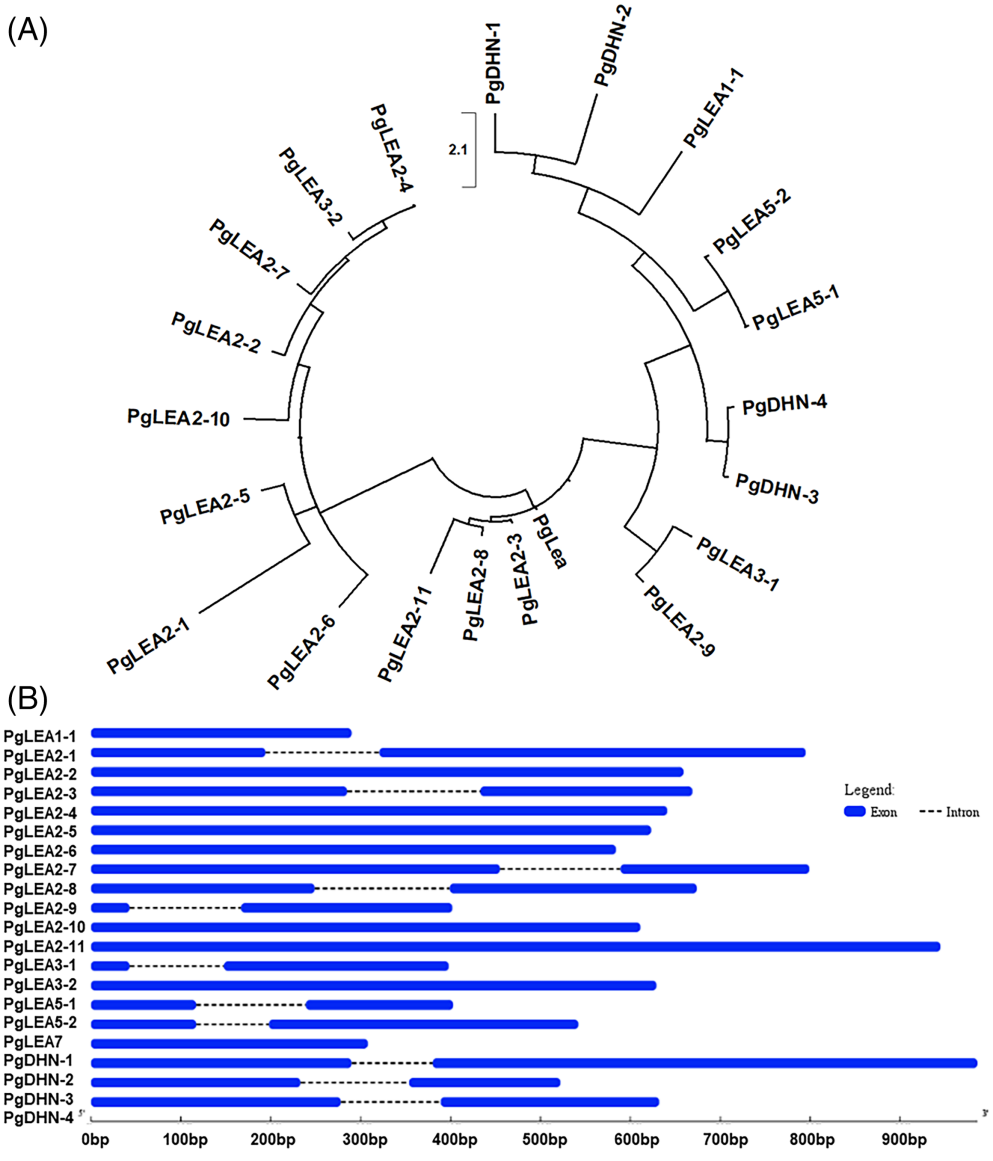

FIGURE 1 Sequence analysis of *PgLEA* family genes. (A) Phylogenetic analysis of *PgLEA* gene family. Different colors distinguish other *LEA* gene family groups. The tree was generated using ClustalW in MEGA6 using the full-length amino acid sequences of the respective *PgLEA* proteins. (B) Intron-exon structural organization in *PgLEA* genes. Dotted lines indicate introns and the rounded rectangle represents exons. The intron and exon length can be identified with the scale at the bottom

3 | RESULTS

3.1 | Characterization of the *PgLEA* gene family

Based on the sequence homology with the *LEA* genes from *Arabidopsis*, sorghum, and rice, 21 *PgLEA* genes were identified and classified into *PgLEA1*, *PgLEA2*, *PgLEA3*, *PgLEA5*, *PgLEA7*, and *PgDHN* (Figure 1A; Table 1). To determine the similarity and homology of the *PgLEA* genes, sequence alignments were performed, and an unrooted phylogenetic tree constructed with other *LEAs* taken from related species (Figure S1). The phylogeny showed that there were five paralogs and 17 orthologs that were closely related to *SbLEAs* (Tables 2 and 3). Families such as *PgLEA1*, *PgLEA2*, *PgLEA3*, *PgLEA5*, *PgLEA7*, and *PgDHN* include 1, 11, 2, 2, 1, and 4 members, respectively, with open reading frames (ORFs) ranging between 291 and 945 bp for *PgLEA1-1* and *PgLEA2-11*. Gene structure analysis of *PgLEAs* revealed that all contain one or two exons (Figure 1B) and the genes were distributed on all seven chromosomes. However, chromosome number 1 has been found to be a hotspot with eight *PgLEA* genes that were localized in the nucleus, mitochondria, chloroplast, and cytoplasm (Table 1). While PgLEA1-1 contains the least number of amino acids (96 aa), PgLEA2-11 consists of the highest (314 aa). PgLEA2-9 has a molecular weight of 9704.16 and PgLEA2-11 has 35109.9 Da. Interestingly, all LEA-6 family members have relatively low molecular masses (Table 1). While the PgLEA proteins mostly has high isoelectric points pI >7, a few have low pI <7. These are hydrophilic with a GRAVY value <0. Domain or motif structure analysis revealed 10 conserved motifs, but none was present across the group as indicated by Pfam codes and WebLogo (Figure S2).

3.2 | Quantitative expression of the *PgLEA* gene family in different tissues under stress treatments

To better understand the tissue specific and relative expression pattern of the *LEA* family genes, qRT-PCR was performed using leaf and root tissues in high resolution cross (HRC) genotypes, that is, VPD-insensitive genotype (ICMR 1122) and VPD-sensitive genotype (ICMR 1152) in response to low (control) and high VPD. Tissue specific expression data depicts significantly high level of expression in leaves in comparison with roots, in both the genotypes. While *PgLEA2-1*, *PgLEA2-8*, *PgLEA2-9*, *PgLEA2-11*, *PgLEA3-1*, *PgDHN-1*, *PgDHN-3*, and *PgDHN-4* exhibited higher expression levels under normal conditions, *PgLEA2-4*, *PgLEA2-7*, *PgLEA2-10* showed better expression levels in the roots (Figure 2A). Expressions of the genes *PgLEA2-1*, *PgLEA2-3*, *PgLEA2-4*, *PgLEA2-6*, *PgLEA2-7*, *PgLEA2-9*, *PgLEA2-10*, *PgLEA5-1*, and *PgDHN-4* were high in both leaves and roots (Figure 2A). Under high VPD stress, superior expression levels were recorded in roots than in leaves when the *PgLEA2-1*, *PgLEA2-4*, *PgLEA2-5*, *PgLEA5-2*, *PgDHN-1*, *PgDHN-3*, and *PgDHN-4* genes were assessed where *PgLEA2-2* showed a distinct expression pattern in leaf compared to other *LEA* genes (Figure 2B). Further, *PgLEA2-2*, *PgLEA2-3*, *PgLEA2-5*, *PgLEA5-2*, *PgLEA7*, and *PgDHN-1* were expressed under all the tested condition. Among these genes, *PgLEA-7* showed significant expression under high VPD, especially in root tissues. From the comparative expression data, it was obvious that while the genotype ICMR1152 was markedly influenced by high VPD conditions, the genotype ICMR1122 did not show much variation (Figure 2B).

3.3 | *PgLEAPC*, *PgLpd1*, and *PgLpd2* promoter cloning and sequence analysis

Based on the expression under different abiotic stresses and high VPD, as observed previously, the *PgLEA-7* gene seemed ideally suited for adaptation to stress environments. Hence, the *PgLEA-7* gene promoter (758 bp) belonging to the group 7 LEA was selected for this study. Considering the type of motifs and position, three different promoter constructs (one full length and two truncated) were developed (Figure 3). *Cis*-motifs such as TATA-box, ABRE and motif IIb (involved in ABA-responsiveness), CAAT-box (*cis*-acting element in promoter and enhancer regions), G-box (associated in light responsiveness), GC-motif (involved in anoxic specific inducibility), MNF1 (light-responsive element) were present in all the three promoter regions (Figure 3). In *PgLEAPC* and *PgLpd1* promoter regions, 5′ UTR Py-rich stretch (confers high transcription levels), LTR (involved in low-temperature responsiveness), O2-site (involved in zein metabolism regulation) and Sp1 and box II (light-responsive elements) were present. A-box (*cis*-acting regulatory element), AuxRR-core (involved in auxin responsiveness), CCAAT-box (MYBHv1 binding site), CCGT

TABLE 2 Nonsynonymous to synonymous substitution ratios of paralogues

*PgLea/PgDhn*gene 1	Chr	PgLea/PgDhn gene 2	Chr	No. nonsynonymous sites (N)	No. synonymous sites (S)	Nonsynonymous substitution rate (d_N)	Synonymous substitution rate (d_S)	d_N/d_S
PgLEA1-1	3	PgDHN-1	2	229	56	2.9939	8.4379	0.3548
PgLEA2-4	1	PgLEA2-6	3	458.6	123.4	12.853	0.1298	99
PgLEA2-9	6	PgLEA3-1	1	210	66	1.2167	50.4881	0.0241
PgLEA5-1	1	PgLEA5-2	1	245.8	33.2	0.2366	1.735	0.1363
PgDHN-3	5	PgDHN-4	6	342	57	14.06	6.5949	2.1333

TABLE 3 Non synonymous to synonymous substitution ratios of *PgLEA* orthologs of pearl millet, *Sorghum*, *Oryza*, maize and *Arabidopsis*

*Pg Lea/PgDhn*gene	Chromosome	Ortholog	Chr.	Number of non-synonymous sites (N)	No. synonymous sites (S)	Non-synonymous substitution rate (dN)	Synonymous substitution rate (dS)	dN/dS
Pgl_GLEAN_10018265	1	Sb09g016830	9	223.2	55.8	3.6397	3.0554	1.1913
Pgl_GLEAN_10012637	3	Sb06g026900	6	233.5	51.5	8.0253	0.0811	99
Pgl_GLEAN_10005440	1	Sb09g018000	9	181.1	55.9	9.2201	25.2594	0.365
Pgl_GLEAN_10028629	6	Os01g21250.1	1	281.5	57.5	0.838	40.882	0.0205
Pgl_GLEAN_10020225	4	Sb05g003631	5	519.2	107.8	0.1971	15.578	0.0127
Pgl_GLEAN_10038309	1	Sb08g003720	8	552.8	86.2	4.096	70.1024	0.0584
Pgl_GLEAN_10026514	3	Sb06g016230	6	549.1	113.9	0.3507	8.5119	0.0412
Pgl_GLEAN_10003471	1	Sb06g033580	6	545.6	111.4	0.0944	6.4625	0.0146
Pgl_GLEAN_10038314	1	Sb08g003690	8	519.8	101.2	2.3875	67.4949	0.0354
Pgl_GLEAN_10001929	7	Sb02g035250	2	482.3	123.7	1.1332	1.3822	0.8199
Pgl_GLEAN_10025348	3	Sb04g032400	4	470.4	111.6	0.3706	19.2193	0.0193
Pgl_GLEAN_10003225	1	Sb09g026230	9	532	134	16.234	0.154	99
Pgl_GLEAN_10022312	7	Sb02g008820	2	746	196	2.2081	54.087	0.0408
Pgl_GLEAN_10021053	6	Os01g43530.1	1	423.9	98.1	4.9139	47.9456	0.1.25
Pgl_GLEAN_10002256	2	Os03g45280.1	3	238.5	37.5	1.6826	1.419	1.1858
Pgl_GLEAN_10025327	3	Sb04g032250	4	721.6	127.4	4.0645	0.0411	99
Pgl_GLEAN_10021551	6	Sb09g018420	9	369.2	86.8	4.6924	0.0474	99

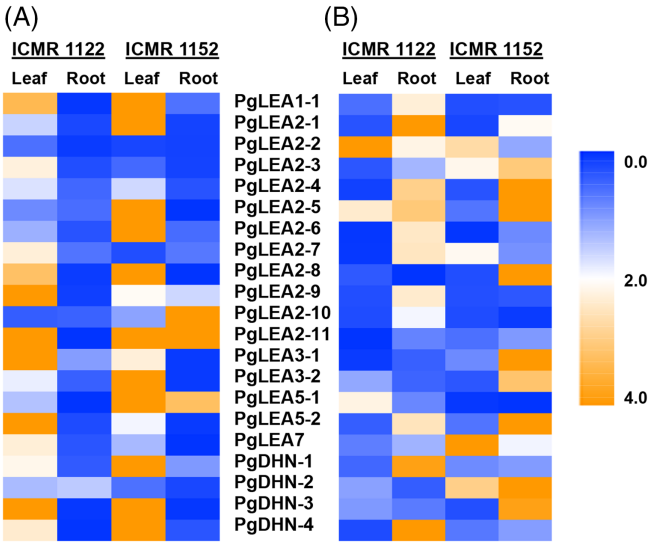

FIGURE 2 Expression analysis of the 21 *PgLEA* genes in leaves and root tissues of *pearl millet* ICMR112 and ICMR 1152 genotypes under control (A) and high VPD (B) conditions. The relative expression of PgLEA genes under high VPD (4.21 kPa) conditions was analyzed using qbase+ software [37], where the data was normalized against two reference genes (PgMDH and PgUBC) and low VPD in leaf and root samples collected at the same time of the day (afternoon). Color scale represents normalized $\log^2$ values where green indicates low expression and red indicates high expression

CC-box (related to meristem specific activation), MBS (MYB binding site implicated in drought-inducibility) and circadian (involved in circadian control) elements were identified (Figure 3; Table S3).

3.4 | Segregation analysis

All the hygromycin-resistant T_1 tobacco seedlings were analyzed for gene integration using PCR and confirmed plants were advanced for T_2 generation and seeds harvested. T_2 seeds of *PgLEAPC*, *PgLEALpd1* and *PgLEALpd2* containing transformants were germinated on MS basal medium supplemented with 50 mg/L hygromycin. These plants showed 100% seed germination and the events were confirmed by PCR and transferred to pots. Three T_2 transgenic events for each promoter were selected for further analysis. Chi-square (χ^2) test was used to find goodness of fit of the transgene segregation with that of Mendelian segregation ratio of 3:1 that was consistent among all of the progeny tested with the presence of a T-DNA insertion locus (Table S2).

3.5 | Induction of *PgLEAPC* promoter regions under control and abiotic stress conditions

To identify the core functional region of the *PgLEAPC* promoter involved under control and abiotic stress responses, a set of serial

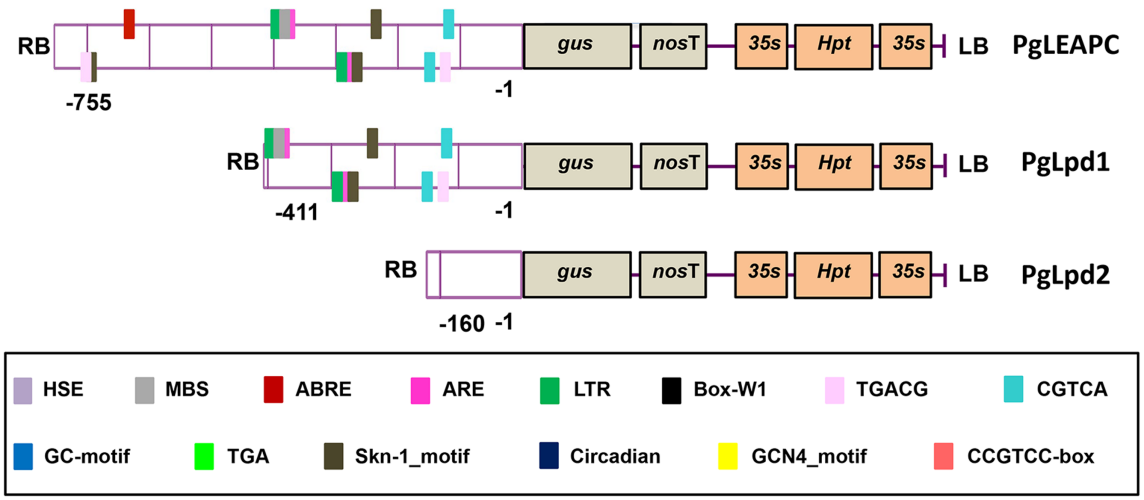

FIGURE 3 *Cis*-motifs position and cloning of the *PgLEA* promoter regions. *Cis*-motifs were identified using the PlantCARE database and cloned in to the plant transformation vector along with *uid*A gene. Gus-*uid*A gene, *nos*T-Nopaline synthase terminator, *Hpt*, Hygromycin phosphotransferase gene; RB, right border; LB, left border

deletions were generated at 406 and 630 bp upstream to the translation starting site. The expression pattern of *PgLEAPC* (758 bp) and its deletions *PgLpd1* (349) and *PgLpd2* (125) were analyzed using GUS histochemical assay (Figure 4). The three promoters displayed spatial expression patterns of *uid*A in histochemical staining under control conditions (Figure 4). While the activity of *PgLEAPC* was mild, activities of *Lpd1* and *Lpd2* were insignificant in the root, a low expression of *uid*A gene was recorded in the leaves of respective transgenics with the three promoters. While *PgLEAPC* promoter showed mild activity, no activity was noticed in the roots of *PgLpd1* and *PgLpd2*. Tobacco transgenic plants harboring the *PgLEAPC* promoter when exposed to heat stress were not significantly affected in comparison with *PgLpd1*, while *PgLpd2* was significantly affected as the wild type controls (data not shown). To check the inducibility of the promoter, tissues exposed to heat stress were stained for *uid*A expression. *PgLEAPC* (Figure 4A) showed the highest expression in the stem and root compared to *PgLpd1* (Figure 4B) which showed mild expression, while the *PgLpd2* (Figure 4C), with the exception for the stem, displayed expression similar to the wild type controls. Under cold stress, the *PgLEAPC* promoter showed high expression of *uid*A in the stem and root and no expression in the leaf (Figure 4A), while *PgLpd1* pro showed less expression in stem and root tissues (Figure 4B) and no expression in *PgLpd2* (Figure 4C). Under salt stress, *PgLEAPC* plants exhibited the highest expression in stem and root tissues in contrast to *PgLpd1* which showed marginal expression in stem and root, and no expression in the leaf. Expression was not noticed in stem, root and leaf tissues of *PgLpd2* similar to the wild type control plants (Figure 4C). When the 30-days-old plants of *PgLEAPC*, *PgLpd1* and *PgLpd2* were exposed to drought stress by withholding water for 5 days, the *uid*A expression was high in the stem and root tissues of *PgLEAPC* plants in contrast to no expression in *PgLpd1* and *PgLpd2* plant tissues.

3.6 | *uid*A gene expression by quantitative real-time PCR

Expression of the *uid*A gene in tobacco transgenics driven by either *PgLEAPC*, *PgLpd1* or *PgLpd2* promoters, under control and different abiotic stress conditions, was further validated using qRT-PCR. Gene expression driven by the promoter *PgLEAPC* was significantly higher in the stem followed by *PgLpd1*, but no expression was observed in *PgLpd2* (Figure 5A). Such a differential expression pattern could be due to the presence of the CCGTCC-box element in the promoter region of *PgLEAPC* which is absent in the other two promoters. Differential expression pattern of *uid*A was high in stem and root tissues of *PgLEAPC* under cold, salt, heat and drought stresses (Figures 5B-E). On the other hand, *PgLpd1* plants showed moderate expression under heat (Figure 5B) and salt stresses in stem and root tissues, and salt stress in the stems of *PgLpd2* (Figure 5D). The *PgLEAPC* promoter was activated under heat, drought, cold, and salt stresses unlike that of *Lpd1* and *Lpd2* which showed no expression. Under salt stress, *PgLEAPC* promoter exhibited the highest *uid*A expression in stem and root tissues while comparatively less expression in the stem and root tissues of *PgLpd1*. In contrast, no *uid*A expression was observed in leaf, root, and stem tissues of *PgLpd2*.

4 | DISCUSSION

LEA proteins are produced in plants in response to dehydration stress where they are implicated in aiding the protection of cytosolic structures, ion sequestration, protein renaturation, transport of nuclear-targeted proteins, prevention of membrane leakage, membrane, and protein stabilization (Chen et al. 2019). *LEA* gene family members have been reported in many crop plants (Liang et al. 2016), especially under

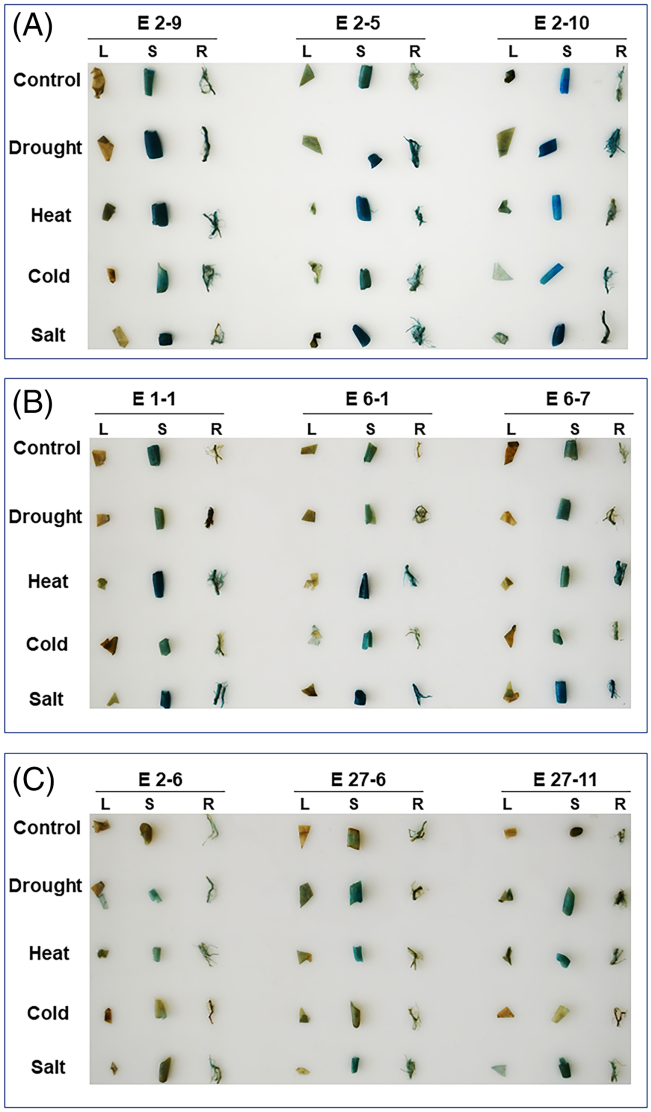

FIGURE 4 Histochemical studies of *uid*A gene controlled by (A) *PgLEAPC*, (B) *PgLpd1*, and (C) *PgLpd2* promoters under different experimental conditions. Transgenic tobacco leaf, stem and roots were treated with different abiotic stress conditions and stained with GUS staining solution and analyzed their activity under non-treated (control) condition and upon treatment with salt (250 mM NaCl for 48 h), heat (4 h), cold (4 h), and drought (4 days). L, leaf; S, stem; R, root

abiotic stresses (Magwanga et al. 2018; Nagaraju et al. 2019). In the present study, 21 *LEA* family genes have been identified in pearl millet, which appear to be less compared with other plants including cereals (Nagaraju et al. 2019; Wang et al. 2007). While some *PgLEA* genes do not have introns, some have a few. These findings are consistent with the earlier studies demonstrated in Brassica (Lan et al. 2013). Transcript production is delayed if introns exist in such genes that may show a negative effect on gene expression. Introns can further extend the nascent transcript's length, adding on additional constraints for transcription (Jeffares et al. 2008). While the number of motifs and composition of each family of LEA2 genes vary, this study detected some amino acid-rich regions like the Gly-rich region in *Arabidopsis*. Although the most conserved motif is rich in lysine (K) residues (Hundertmark & Hincha 2008), the amino acid composition of LEA proteins has been shown to be disordered (Dure III 1989, 1993). While the LEA proteins are relatively small and intrinsically unstructured, they play essential roles in cells, likely by forming flexible residual structural elements (Fuxreiter et al. 2004) such as α-helical structures and polyproline II (PII) helices (Soulages et al. 2003). Even though the *PgLEA*s showed a closer orthologous relationship with *S. bicolor*, only six subgroups were observed in pearl millet in contrast to eight subgroups reported in sorghum (Nagaraju et al. 2019). Like rice (Wang et al. 2007) and tomato (Cao & Li 2015), the absence of LEA6 group in pearl millet could indicate the existence of variation in the LEA groups in some plants. In sorghum (Nagaraju et al. 2019), the hydrophobic LEA2 is reported as a larger group which is in contrast to that of *Arabidopsis* (Hundertmark & Hincha 2008) and *Populus* (Lan et al. 2013). It appears that whole-genome duplication contributes to expansion of the SbLEA family in sorghum but not in rice. Perhaps the natural selection might have influenced the *PgLEA* genes during evolution, similar to *Brassica* (Liang et al. 2016) and cotton (Magwanga et al. 2018). Analysis of *PgLEA* family genes displayed a differential expression pattern under high VPD stress where *PgLEA-7* gene showed significant expression under abiotic stresses (Reddy et al. 2012) and high VPD conditions. This protein appears to have in vitro chaperonic activity, which when overexpressed in bacteria provided protection from the damage caused by salt and heat stresses, thereby indicating its key role under stress conditions (Reddy et al. 2012). When an analysis of the *PgLEA-7* gene and its promoter was carried out in silico, distinct *cis*-elements alongside transcription factor binding sites were also noticed that appeared to be important for imparting stress tolerance (Reddy et al. 2012). Hence, in the present study, two deletions were made in the *PgLEA* promoter viz., *PgLpd1* and *PgLpd2* with lengths of 758, 349 and 125 bp, respectively, for subsequent use.

The expression of dehydration-responsive element binding protein/C-repeat-binding factor (DREBs/CBF) from stress-inducible *rd29* promoter has been shown to minimize not only the negative effects on growth and yield, but also improved stress tolerance in comparison with *CaMV 35S* promoter (Kasuga et al. 1999, 2004). In fact, some promoters get activated by ABA, drought, osmotic and salt stresses (Kim et al. 2002). Moreover, core-, proximal- and long-range regulatory elements constitute the *cis*-regulatory elements and are involved in stress-responsive expression of genes (Yamaguchi-Shinozaki & Shinozaki 2005). It is the RNA polymerase II and its subunits alongside the *trans*-acting factors which recognize the core elements and functional DNA sequences that initiate the process of transcription (Lenka et al. 2019). ABA-responsive elements (ABRE), and dehydration-responsive element/C repeat (DRE/CRT) are the major *cis*-acting elements known in stress-inducible gene expressions (Kim et al. 2002). The core promoter elements also contain the TATA box, which along with the downstream promoter elements help in gene activation. To identify efficient stress-inducible promoters for genetic engineering applications, the promoter region of the stress-inducible gene *PgLEAPC* was functionally characterized. *LEA* genes such as *LEA*3 and *DHN-5* from wheat (Brini et al. 2007; Wang et al. 2008), *HVA1* from barley (Xu et al. 1996), and

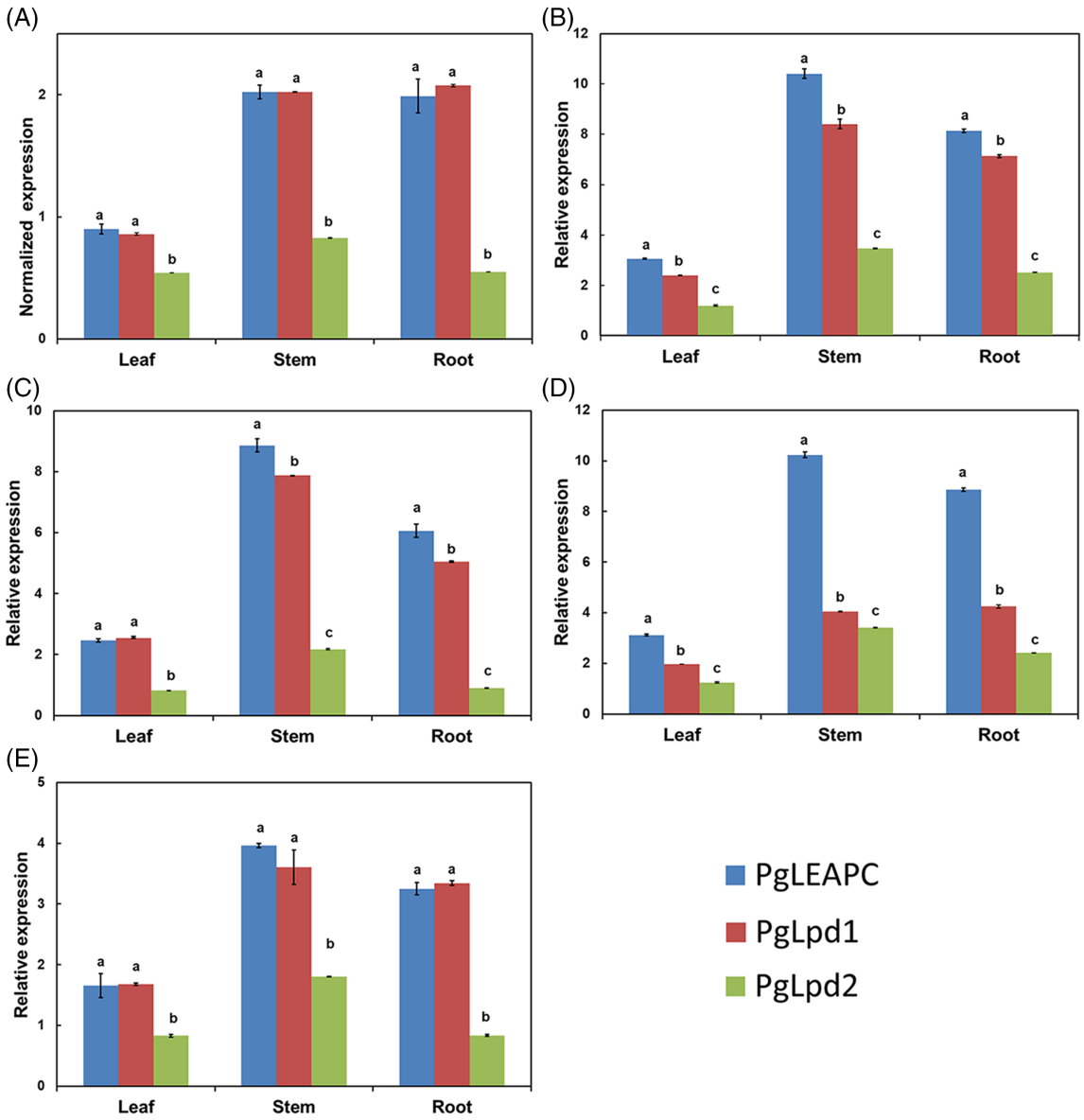

FIGURE 5 Expression analysis of *uid*A transcript in leaf, stem, and root tissues of *PgLEAPC*, *PgLpd1* and *PgLpd2* transgenic tobacco plants under control and abiotic stress conditions. (A) Control, (B) heat stress, (C) cold stress, (D) salt stress, and (E) drought stress. Values are expression folds obtained after normalizing against two reference genes (*NtUBC* and *NtEF-1α*). All samples were analyzed in triplicates in three independent experiments. Bars with different letters are significantly different ($p < 0.05$)

*BhLEA*1 and *BhLEA*2 from the resurrection plant *Boea hygrometrica* have been shown to improve drought tolerance. A *LEA* gene (*IpLEA*) from *Ipomoea pescaprae*, which belongs to group 2 LEAs confers salt stress tolerance in transgenic *Arabidopsis* plants (Zheng et al. 2019). Since a single transgene transformation might not be sufficient to produce abiotic tolerant crops, multiple transgenes if expressed with the same promoters may lead to homology-dependent gene silencing (De Wilde et al. 2000; Han et al. 2015; Verdaguer et al. 1996). Hence, the utilization of heterologous promoters to regulate expression of gene/s of interest in host plants might help to prevent homology-dependent gene silencing (Dong et al. 2015; Kumpatla et al. 1998). For spatio-temporal gene expression, since both *cis*-elements and *trans*-acting factors play pivotal roles, it is now feasible to identify stress-responsive elements and the corresponding regulatory genes for manipulation utilizing *Agrobacterium*-mediated transformations. However, a proper characterization of stress-inducible promoters is crucial for their subsequent use in genetic engineering. Several promoters inducible by abiotic stresses have been isolated from *Atriplex centralasiatica* (Yin et al. 2002), *Thellungiella halophila* (Sun et al. 2010), *Ammopiptanthus mongolicust* (Guo et al. 2010) and *Salicornia brachiata* (Tiwari et al. 2016) and have been found to have stress-responsive *cis*-regulatory motifs. Rerksiri et al. (2013) identified six heat-responsive genes in rice which also responded to ABA, polyethylene glycol, and salt, but not cold treatment. Three of the six promoters (*OsHsfB2cp*, *PM19p*, and *Hsp90p*) displayed very high heat-induced *uid*A expression and moderate drought-induced activities. Reconstruction of *cis*-elements in these rice promoters lead to

the development of efficient heat-inducible promoters for genetic engineering studies. A novel *LmSAP* gene promoter in *Lobularia maritima* showed stress responsiveness, thereby suggesting that it can be an excellent tool for stress-inducible gene expression and genetic engineering (Saad et al. 2020). Hou et al. (2012) constructed three synthetic promoters including EKCM, EKCRM and ECCRM with multiple *cis*-acting, stress-responsive elements that were fused separately with the promoters with the *uid*A gene for expression studies in *A. thaliana*. It was observed that the higher expression of *uid*A gene activity was driven by the three synthetic promoters under cold, salt, and desiccation stresses, besides a higher *uid*A activity in comparison with the *rd29* promoter. Our results also suggest that *PgLEAPC* promoter could be used to confer high levels of gene activity under drought, heat, and salt stresses in transgenics. Further, it infers that monocot promoter can be an ideal candidate for conferring drought and heat tolerance in dicotyledonous crops. Since, the *PgLEAPC* promoter isolated from pearl millet shows high transcriptional activity during drought and heat stress conditions and is inducible by multiple abiotic stresses, the promoter can be used for obtaining superior crops with better tolerance to harsh environmental conditions. While the *PgLEAPC* promoter displayed better *uid*A expression under a wide range of abiotic stresses, the remaining two *PgLpd1* and *PgLpd2* truncated promoters could not exhibit such an activity possibly due to the lack of stress inducible *cis*-elements like MBS and CCAAT-box that are crucial for responses to abiotic stresses. Thus, the stress-inducible promoters represent highly valuable tools that can effective be deployed for generating transgenics plants for tolerance to abiotic stresses.

5 | CONCLUSIONS

In the present study, a total 21 *PgLEA* genes were identified and systematically characterized. Besides, the *PgLEAPC* promoter was functionally characterized by making the truncated versions of the promoter region and analyzed their activity under heat, cold, salt, and drought stress conditions. A novel *PgLEAPC* promoter from pearl millet was functionally validated which may be of immense value for regulating the genes under drought, salt, cold and heat stresses. Such promoters could be used in transgenic breeding as allele-specific markers for developing heat and drought stress tolerant crops depending on the heterologous promoter functionality.

ACKNOWLEDGMENTS

Palakolanu Sudhakar Reddy acknowledges the Department of Science and Technology, Government of India for the research grant through the INSPIRE Faculty Award No. IFALSPA-06 and the partial support from the Core Grant Research (CRG) grant file no. DST No: EMR/20l6/006726. Thanks to the Dr. Rajeev Varshney and Mr. Prasad Bajaj, SISU theme, ICRISAT, for the sequence extraction from the pearl millet genome. This study was undertaken as part of the CGIAR Research Program on Grain Legumes & Dryland Cereals (CRP-GLDC). ICRISAT is a CGIAR research center.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Palakolanu Sudhakar Reddy, Pooja Bhatnagar Mathur and Kiran K. Sharma designed the experiments, Palakolanu Sudhakar Reddy, Polavarapu Kavi Kishor and Kummari Divya executed the study, Palakolanu Sudhakar Reddy, Vincent Vadez, Aishwarya Shankhapal Rajesh, Polavarapu Kavi Kishor and Kummari Divya analyzed data. Kummari Divya, Palakolanu Sudhakar Reddy, Kiran K. Sharma, Polavarapu Kavi Kishor and Pooja Bhatnagar Mathur designed and wrote the manuscript and refined.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon request and data supporting the findings of this study are available in the supplementary material of this article.

ORCID

Sudhakar Reddy Palakolanu https://orcid.org/0000-0002-5341-187X

Polavarapu Kavi Kishor https://orcid.org/0000-0002-9587-7547

REFERENCES

Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L. & Noble, W.S. (2009) MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Research*, 37 (suppl-2), W202-W208.

Banerjee, J., Sahoo, D.K., Dey, N., Houtz, R.L. & Maiti, I.B. (2013) An intergenic region shared by At4g35985 and At4g35987 in *Arabidopsis thaliana* is a tissue specific and stress inducible bidirectional promoter analyzed in transgenic Arabidopsis and tobacco plants. *PLoS One*, 8(11), e79622.

Battaglia, M., Olvera-Carrillo, Y., Garciarrubio, A., Campos, F. & Covarrubias, A.A. (2008) The enigmatic LEA proteins and other hydrophilins. *Plant Physiology*, 148, 6–24.

Bhatnagar-Mathur, P., Devi, J., Lavanya, M., Vadez, V., Serraj, R., Yamaguchi-Shinozaki, K. et al. (2007) Stress-inducible expression of *At DREB1A* in transgenic peanut (*Arachis hypogaea* L.) increases transpiration efficiency under water-limiting conditions. *Plant Cell Reports*, 26, 2071–2082.

Bhatnagar-Mathur, P., Vadez, V. & Sharma, K.K. (2008) Transgenic approaches for abiotic stress tolerance in plants: retrospect and prospects. *Plant Cell Reports*, 27, 411–424.

Brini, F., Hanin, M., Mezghani, I., Berkowitz, G.A. & Masmoudi, K. (2007) Overexpression of wheat Na^+/H^+ antiporter TNHX1 and H^+-pyrophosphatase TVP1 improve salt-and drought-stress tolerance in *Arabidopsis thaliana* plants. *Journal of Experimental Botany*, 58, 301–308.

Cao, J. & Li, X. (2015) Identification and phylogenetic analysis of late embryogenesis abundant proteins family in tomato (*Solanum lycopersicum*). *Planta*, 241, 757–772.

Charfeddine, M., Saïdi, M.N., Charfeddine, S., Hammami, A. & Bouzid, R.G. (2015) Genome-wide analysis and expression profiling of the ERF transcription factor family in potato (*Solanum tuberosum* L.). *Molecular Biotechnology*, 57, 348–358.

Chen, Y., Li, C., Zhang, B., Yi, J., Yang, Y., Kong, C. et al. (2019) The role of the late embryogenesis-abundant (LEA) protein family in development and the abiotic stress response: a comprehensive expression analysis of potato (*Solanum tuberosum*). *Genes*, 10(2), 148.

Dalal, M., Tayal, D., Chinnusamy, V. & Bansal, K.C. (2009) Abiotic stress and ABA-inducible group 4 LEA from *Brassica napus* plays a key role in salt and drought tolerance. *Journal of Biotechnology*, 139, 137–145.

De Wilde, C., Van Houdt, H., De Buck, S., Angenon, G., De Jaeger, G. & Depicker, A. (2000) Plants as bioreactors for protein production: avoiding the problem of transgene silencing. *Plant Molecular Biology*, 43, 347–359.

Divya, K., Kishor, P.K., Bhatnagar-Mathur, P., Singam, P., Sharma, K. K., Vadez, V., et al. (2019) Isolation and functional characterization of three abiotic stress-inducible (Apx, Dhn and Hsc70) promoters from pearl millet (Pennisetum glaucum L.). *Molecular biology reports*, 46(6), 6039–6052.

Divya, K., Reddy, P.S., Kavi Kishor, P.B., Bhatnagar-Mathur, P., Singam, P., Vadez, V. et al. (2020) An update and perspectives on the use of promoters in plant genetic engineering. *Journal of Biosciences*, 45, 119. https://doi.org/10.1007/s12038-020-00087-6

Dong, P., Xiong, F., Que, Y., Wang, K., Yu, L., Li, Z. et al. (2015) Expression profiling and functional analysis reveals that TOR is a key player in regulating photosynthesis and phytohormone signaling pathways in *Arabidopsis*. *Frontiers in Plant Science*, 7(6), 677.

Du, D., Zhang, Q., Cheng, T., Pan, H., Yang, W. & Sun, L. (2013) Genome-wide identification and analysis of late embryogenesis abundant (LEA) genes in *Prunus mume*. *Molecular Biology Reports*, 40, 1937–1946.

Du, H., Shen, X., Huang, Y., Huang, M. & Zhang, Z. (2016) Overexpression of *Vitreoscilla* hemoglobin increases waterlogging tolerance in *Arabidopsis* and maize. *BMC Plant Biology*, 16, 35–35.

Dure, L., III. (1989) Characteristics of the storage proteins of cotton. *Journal of the American Oil Chemists Society*, 66, 356–359.

Dure, L., III. (1993) Structural motifs in Lea proteins. In: Close, T.J. & Bray, E.A. (Eds.) *Plant responses to cellular dehydration during environmental stress. Current topics in plant physiology*, Vol. 10. Rockville, MD: American Society of Plant Physiologists Series, pp. 91–103.

Federspiel, N. (2000) Deciphering a weed. Genomic sequencing of *Arabidopsis*. *Plant Physiology*, 124, 1456–1459.

Filiz, E., Ozyigit, I.I., Tombuloglu, H. & Koc, I. (2013) In silico comparative analysis of LEA (late embryogenesis abundant) proteins in *Brachypodium distachyon* L. *Plant Omics*, 6, 433–440.

Fuxreiter, M., Simon, I., Friedrich, P. & Tompa, P. (2004) Preformed structural elements feature in partner recognition by intrinsically unstructured proteins. *Journal of Molecular Biology*, 338, 1015–1026.

Guo, L., Yu, Y., Xia, X. & Yin, W. (2010) Identification and functional characterisation of the promoter of the calcium sensor gene CBL1 from the xerophyte *Ammopiptanthus mongolicus*. *BMC Plant Biology*, 10, 18.

Han, Y.J., Kim, Y.M., Hwang, O.J. & Kim, J.I. (2015) Characterization of a small constitutive promoter from *Arabidopsis* translationally controlled tumor protein (AtTCTP) gene for plant transformation. *Plant Cell Reports*, 34, 265–275.

Hellemans, J., Mortier, G., De Paepe, A., Speleman, F. & Vandesompele, J. (2007) qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biology*, 8, 1–14.

Higo, K., Ugawa, Y., Iwamoto, M. & Korenaga, T. (1999) Plant *cis*-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Research*, 27, 297–300.

Horton, P., Park, K.J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C. J. & Nakai, K. (2007) WoLF PSORT: protein localization predictor. *Nucleic Acids Research*, 35 (suppl_2), W585-W587.

Hou, L., Chen, L., Wang, J., Xu, D., Dai, L., Zhang, H. et al. (2012) Construction of stress responsive synthetic promoters and analysis of their activity in transgenic *Arabidopsis thaliana*. *Plant Molecular Biology Reporter*, 30, 1496–1506.

Hundertmark, M. & Hincha, D.K. (2008) LEA (late embryogenesis abundant) proteins and their encoding genes in *Arabidopsis thaliana*. *BMC Genomics*, 9, 118.

Jeffares, D.C., Penkett, C.J. & Bähler, J. (2008) Rapidly regulated genes are intron poor. *Trends in Genetics*, 24, 375–378.

Jefferson, R.A., Kavanagh, T.A. & Bevan, M.W. (1987) GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO Journal*, 6, 3901–3907.

Jin, B., Sheng, Z., Muhammad, I., Chen, J. & Yang, H. (2019) Cloning and functional analysis of the promoter of a stress-inducible gene (*Zmap*) in maize. *PLoS One*, 14, e0211941. https://doi.org/10.1371/journal.pone.0211941

Kamisugi, Y. & Cuming, A.C. (2005) The evolution of the abscisic acid-response in land plants: comparative analysis of group 1 LEA gene expression in moss and cereals. *Plant Molecular Biology*, 59, 723–737.

Kasuga, M., Liu, Q., Miura, S., Yamaguchi-Shinozaki, K. & Shinozaki, K. (1999) Improving plant drought, salt, and freezing tolerance by gene transfer of a single stress-inducible transcription factor. *Nature Biotechnology*, 17, 287–291.

Kasuga, M., Miura, S., Shinozaki, K. & Yamaguchi-Shinozaki, K. (2004) A combination of the Arabidopsis DREB1A gene and stress-inducible rd29A promoter improved drought- and low-temperature stress tolerance in tobacco by gene transfer. *Plant Cell Physiology*, 45, 346–350.

Kholova, J. & Vadez, V. (2013) Water extraction under terminal drought explains the genotypic differences in yield, not the antioxidant changes in leaves of pearl millet (*Pennisetum glaucum*). *Functional Plant Biology*, 40, 44–53.

Kim, H.J., Kim, Y.K., Park, J.Y. & Kim, J. (2002) Light signaling mediated by phytochrome plays an important role in cold-induced gene expression through the C-repeat/dehydration responsive element (C/DRE) in *Arabidopsis thaliana*. *The Plant Journal*, 29, 693–704.

Kim, I.J., Lee, J., Han, J.A., Kim, C.S. & Hur, Y. (2011) Citrus lea promoter confers fruit-preferential and stress inducible gene expression in *Arabidopsis*. *Canadian Journal of Plant Science*, 91, 459–466.

Kishor, P.B.K., Renuka, S., Rajasheker, G., Nagaraju, M., Shridhar, K., Divya, D. et al. (2020) Lysine, lysine-rich, serine and serine-rich proteins: link between metabolism, development and abiotic stress tolerance and the role of ncRNAs in their regulation. *Frontiers in Plant Science*, 11, 546213. https://doi.org/10.3389/fpls.2020.546213

Kumpatla, S.P., Chandrasekharan, M.B., Iyer, L.M., Guofu, L. & Hall, T.C. (1998) Genome intruder scanning and modulation systems and transgene silencing. *Trends in Plant Science*, 3, 97–104.

Lan, T., Gao, J. & Zeng, Q.Y. (2013) Genome-wide analysis of the LEA (late embryogenesis abundant) protein gene family in *Populus trichocarpa*. *Tree Genetics Genomes*, 9, 253–264.

Lenka, S.K., Singh, A.K., Muthusamy, S.K., Smita, S., Chinnusamy, V. & Bansal, K.C. (2019) Heterologous expression of rice RNA-binding glycine-rich (RBG) gene OsRBGD3 in transgenic *Arabidopsis thaliana* confers cold stress tolerance. *Functional Plant Biology*, 46(5), 482–491.

Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y. et al. (2002) PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Research*, 30, 325–327.

Liang, Y., Xiong, Z., Zheng, J., Xu, D., Zhu, Z., Xiang, J. et al. (2016) Genome-wide identification, structural analysis and new insights into late embryogenesis abundant (LEA) gene family formation pattern in *Brassica napus*. *Scientific Reports*, 6, 24265.

Liu, H., Xing, M., Yang, W., Mu, X., Wang, X., Lu, F. et al. (2019) Genome-wide identification of and functional insights into the late embryogenesis abundant (LEA) gene family in bread wheat (*Triticum aestivum*). *Scientific Reports*, 9, 1–11.

Livak, K.J. & Schmittgen, T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2-\Delta\Delta CT$ method. *Methods*, 25, 402–408.

Magwanga, R.O., Lu, P., Kirungu, J.N., Lu, H., Wang, X., Cai, X. et al. (2018) Characterization of the late embryogenesis abundant (LEA) proteins family and their role in drought stress tolerance in upland cotton. *BMC Genetics*, 19, 6.

Malnoy, M., Reynoird, J.P., Borejsza-Wysocka, E.E. & Aldwinckle, H.S. (2006) Activation of the pathogen-inducibleGst1 promoter of potato

after elicitation by *Venturia inaequalis* and *Erwinia amylovora* in transgenic apple (*Malus* X *domestica*). *Transgenic Research*, 15, 83–93.

Mehrotra, R., Gupta, G., Sethi, R., Bhalothia, P., Kumar, N. & Mehrotra, S. (2011) Designer promoter: an artwork of cis engineering. *Plant Molecular Biology*, 75, 527–536.

Murashige, T. & Skoog, F. (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiologia Plantarum*, 15, 473–497.

Nagaraju, M., Kumar, S.A., Reddy, P.S., Kumar, A., Rao, D.M. & Kavi Kishor, P.B. (2019) Genome-scale identification, classification, and tissue specific expression analysis of late embryogenesis abundant (LEA) genes under abiotic stress conditions in *Sorghum bicolor* L. *PLoS One*, 14, e0209980.

Park, S.C., Kim, Y.H., Jeong, J.C., Kim, C.Y., Lee, H.S., Bang, J.W. et al. (2011) Sweet potato late embryogenesis abundant 14 (IbLEA14) gene influences lignification and increases osmotic-and salt stress-tolerance of transgenic calli. *Planta*, 233, 621–634.

Pearson, K. (1900) On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 50, 157–175.

Reddy, P.S., Jogeswar, G., Rasineni, G.K., Maheswari, M., Reddy, A.R., Varshney, R.K. et al. (2015a) Proline over-accumulation alleviates salt stress and protects photosynthetic and antioxidant enzyme activities in transgenic sorghum [*Sorghum bicolor* (L.) Moench]. *Plant Physiology and Biochemistry*, 94, 104–113.

Reddy, P.S., Reddy, D.S., Sharma, K.K., Bhatnagar-Mathur, P. & Vadez, V. (2015b) Cloning and validation of reference genes for normalization of gene expression studies in pearl millet [*Pennisetum glaucum* (L.) R. Br.] by quantitative real-time PCR. *Plant Gene*, 1, 35–42.

Reddy, P.S., Reddy, G.M., Pandey, P., Chandrasekhar, K. & Reddy, M.K. (2012) Cloning and molecular characterization of a gene encoding late embryogenesis abundant protein from *Pennisetum glaucum*: protection against abiotic stresses. *Molecular Biology Reports*, 39, 7163–7174.

Reddy, P.S., Tharanya, M., Sivasakthi, K., Srikanth, M., Hash, C.T., Kholova, J. et al. (2017) Molecular cloning and expression analysis of aquaporin genes in pearl millet [*Pennisetum glaucum* (L) R. Br.] genotypes contrasting in their transpiration response to high vapour pressure deficits. *Plant Science*, 265, 167–176.

Rerksiri, W., Zhang, X., Xiong, H. & Chen, X. (2013) Expression and promoter analysis of six heat stress-inducible genes in rice. *The Scientific World Journal*, 397401, 1–9. https://doi.org/10.1155/2013/397401

Saad, R.B., Romdhane, W.B., Zouari, N., Hsouna, A.B., Harbaoui, M. Brini, F, & Ghneim-Herrera, T. (2020) Characterization of a novel *LmSAP* gene promoter from *Lobularia maritima*: tissue specificity and environmental stress responsiveness. *PLoS One*, 15. e0236943. https://doi.org/10.1371/journal.pone.0236943

Schmidt, G.W. & Delaney, S.K. (2010) Stable internal reference genes for normalization of real-time RT-PCR in tobacco (*Nicotiana tabacum*) during development and abiotic stress. *Molecular Genetics and Genomics*, 283, 233–241.

Soltis, D.E. & Soltis, P.S. (2003) The role of phylogenetics in comparative genetics. *Plant Physiology*, 132, 1790–1800.

Soulages, J.L., Kim, K., Arrese, E.L., Walters, C. & Cushman, J.C. (2003) Conformation of a group 2 late embryogenesis abundant protein from soybean. Evidence of poly (L-proline)-type II structure. *Plant Physiology*, 131, 963–975.

Sun, Q., Gao, F., Zhao, L., Li, K. & Zhang, J. (2010) Identification of a new 130 bp *cis*-acting element in the TsVP1 promoter involved in the salt stress response from *Thellungiella halophila*. *BMC Plant Biology*, 10, 90. https://doi.org/10.1186/1471-2229-10-90

Tiwari, V., Patel, M.K., Chaturvedi, A.K., Mishra, A. & Jha, B. (2016) Functional characterization of the tau class glutathione-S-transferases gene (SbGSTU) promoter of *Salicornia brachiata* under salinity and osmotic stress. *PLoS One*, 11, e0148494. https://doi.org/10.1371/journal.pone.0148494

Vadez, V., Kholova, J., Medina, S., Kakkera, A. & Anderberg, H. (2014) Transpiration efficiency: new insights into an old story. *Journal of Experimental Botany*, 65, 6141–6153.

Varshney, R.K., Shi, C., Thudi, M., Mariac, C., Wallace, J., Qi, P. et al. (2017) Pearl millet genome sequence provides a resource to improve agronomic traits in arid environments. *Nature Biotechnology*, 35(10), 969–976.

Verdaguer, B., de Kochko, A., Beachy, R.N. & Fauquet, C. (1996) Isolation and expression in transgenic tobacco and rice plants, of the cassava vein mosaic virus (CVMV) promoter. *Plant Molecular Biology*, 31, 1129–1139.

Wang, B.F., Wang, Y.C., Zhang, D.W., Li, H.Y. & Yang, C.P. (2008) Verification of the resistance of a LEA gene from *Tamarix* expression in *Saccharomyces cerevisiae* to abiotic stresses. *Journal of Forestry Research*, 19, 58–62.

Wang, G.T., Hu, Z.H. & Chen, J.Q. (2003) CBF transcriptional activators and their roles in enhancing stress tolerance. *Plant Physiology Communications*, 39, 402–410.

Wang, M., Li, P., Li, C., Pan, Y., Jiang, X., Zhu, D. et al. (2014) SiLEA14, a novel atypical LEA protein, confers abiotic stress resistance in foxtail millet. *BMC Plant Biology*, 14, 290.

Wang, X.S., Zhu, H.B., Jin, G.L., Liu, H.L., Wu, W.R. & Zhu, J. (2007) Genome-scale identification and analysis of LEA genes in rice (*Oryza sativa* L.). *Plant Science*, 172, 414–420.

Xiao, B., Huang, Y., Tang, N. & Xiong, L. (2007) Over-expression of a LEA gene in rice improves drought resistance under the field conditions. *Theoretical and Applied Genetics*, 115, 35–46.

Xu, D., Duan, X., Wang, B., Hong, B., Ho, T.H.D. & Wu, R. (1996) Expression of a late embryogenesis abundant protein gene, HVA1, from barley confers tolerance to water deficit and salt stress in transgenic rice. *Plant Physiology*, 110, 249–257.

Yamaguchi-Shinozaki, K. & Shinozaki, K. (2005) Organization of *cis*-acting regulatory elements in osmotic- and cold-stress-responsive promoters. *Trends in Plant Science*, 10, 88–94.

Yan, Y., Jia, H., Wang, F., Wang, C., Liu, S. & Guo, X. (2015) Overexpression of GhWRKY27a reduces tolerance to drought stress and resistance to *Rhizoctonia solani* infection in transgenic *Nicotiana benthamiana*. *Frontiers in Physiology*, 6, 265.

Yin, Y., Wang, Z.Y., Mora-Garcia, S., Li, J., Yoshida, S., Asami, T. et al. (2002) BES1 accumulates in the nucleus in response to brassinosteroids to regulate gene expression and promote stem elongation. *Cell*, 109, 181–191.

Zhang, J.F., Deng, X.P. & Mu, X.Q. (2002) Plant aquaporin. *Plant Physiology Communications*, 38, 88–91.

Zhang, X., Lu, S., Jiang, C., Wang, Y., Lv, B., Shen, J. et al. (2014) RcLEA, a late embryogenesis abundant protein gene isolated from *Rosa chinensis*, confers tolerance to *Escherichia coli* and *Arabidopsis thaliana* and stabilizes enzyme activity under diverse stresses. *Plant Molecular Biology*, 85, 333–347.

Zheng, J., Su, H., Lin, R., Zhang, H., Xia, K., Jian, S. et al. (2019) Isolation and characterization of an atypical LEA gene (IpLEA) from *ipomoea pes-caprae* conferring salt/drought and oxidative stress tolerance. *Scientific Reports*, 9, 1–21.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Divya, K., Palakolanu, S.R., Kavi Kishor, P., Rajesh, A.S., Vadez, V., Sharma, K.K. et al. (2021) Functional characterization of late embryogenesis abundant genes and promoters in pearl millet (*Pennisetum glaucum* L.) for abiotic stress tolerance. *Physiologia Plantarum*, 1–13. Available from: https://doi.org/10.1111/ppl.13544