Plant Physiology and Biochemistry 135 (2019) 277–286

Contents lists available at ScienceDirect

Plant Physiology and Biochemistry

journal homepage: www.elsevier.com/locate/plaphy

Research article

The late embryogenesis abundant gene family in tea plant (*Camellia sinensis*): Genome-wide characterization and expression analysis in response to cold and dehydration stress

Weidong Wang, Tong Gao, Jiangfei Chen, Jiankun Yang, Huiyu Huang, Youben Yu*

College of Horticulture, Northwest A&F University, Yangling, 712100, Shaanxi, China

ARTICLE INFO

Keywords:
Cold stress
Dehydration stress
Expression pattern
LEA
Tea plant

ABSTRACT

Late embryogenesis abundant (LEA) proteins are a large and highly diverse family of polypeptides that play important roles in plant growth, development and stress responses. At present, *LEA* gene families have been identified and systematically characterized in many plant species. However, the *LEA* gene family in tea plant has not been revealed, and the biological functions of the members of this family remain unknown. In this study, 33 *CsLEA* genes were identified from tea plant via a genome-wide study, and they were clustered into seven groups according to analyses of their phylogenetic relationships, gene structures and protein conserved motifs. In addition, expression analysis revealed that the *CsLEA* genes were specifically expressed in one or more tissues and significantly induced under cold and dehydration stresses, implying that *CsLEA* genes play important roles in tea plant growth, development and response to cold and dehydration stresses. Furthermore, a potential transcriptional regulatory network, including DREB/CBF, MYB, bZIP, bHLH, BPC and other transcription factors, is directly associated with the expression of *CsLEA* genes, which may be ubiquitous and important in the above mentioned processes. This study could help to increase our understanding of CsLEA proteins and their contributions to stress tolerance in tea plant.

1. Introduction

Abiotic stress factors, such as cold, heat, drought, salinity and heavy metals, severely limit plant growth and development as well as the final yield and quality of crops. Generally, plants have developed a series of defence mechanisms to respond to these stress factors (Kosova et al., 2014). An important group of highly hydrophilic proteins, called late embryogenesis abundant (LEA) proteins, accumulate naturally for cellular protection during abiotic stress tolerance (Zhao et al., 2011). LEA proteins were first described to accumulate to high amounts in the late stages of cotton seed development (Dure et al., 1981), and they were later identified in seeds of other plant species and in vegetative plant organs (Hincha and Thalhammer, 2012). However, LEA proteins are not plant specific; they have also been identified in cyanobacteria, bacteria and some invertebrates (Browne et al., 2002; Close and Lammers, 1993; Hatanaka et al., 2015; Stacy and Aalen, 1998). Importantly, their presence has frequently been associated with cellular tolerance, which may be induced by abiotic stresses, particularly water deficit and cold stress (Salleh et al., 2012). Under extreme stress, LEA proteins, as molecular chaperones, protect against protein aggregation because of their characteristics that include being small in size, being highly hydrophilic and having low complexity (Shih et al., 2008).

In plants, LEA proteins are divided into at least eight different groups according to their specific domains and distinct motif compositions, including LEA_1, LEA_2, LEA_3, LEA_4, LEA_5, LEA_6, dehydrin and seed maturation protein (SMP) groups, but the groupings and the nomenclature of the groups are not consistent in the literature (Hincha and Thalhammer, 2012). Many studies have characterized the functions of LEA proteins, especially their roles in stress responses. For example, OsEm1, a LEA_1 group protein in rice, can enhance the tolerance of rice to drought (Yu et al., 2016). Overexpression of *Chimonanthus praecox CpLEA5* in Arabidopsis provided enhanced osmotic and low-temperature tolerance (Liu et al., 2015a). Similarly, the expression of *ZmLEA3*, *AtLEA5*, *ShDHN* and *SiLEA14*, etc. could enhance tolerance to different stresses in transgenic plants (Liu et al., 2015b, 2016; Salleh et al., 2012; Wang et al., 2014). These studies indicated that the accumulation of LEA proteins is critical for plant resistance. In addition, the functional mechanisms of LEA proteins associated with stress tolerance have been reported, including protection of cellular structures from the effects of water loss and desiccation, protection of proteins from stress-induced

---

* Corresponding author.
E-mail address: yyben@163.com (Y. Yu).

https://doi.org/10.1016/j.plaphy.2018.12.009
Received 8 September 2018; Received in revised form 14 December 2018; Accepted 15 December 2018

damage, sequestration of ions, and folding of denatured proteins (Hand et al., 2011; Shao et al., 2005). At present, *LEA* gene families have been comprehensively analysed in various plant species, such as Arabidopsis, rice, poplar, maize and tomato (Cao and Li, 2015; Hundertmark and Hincha, 2008; Lan et al., 2013; Li and Cao, 2016; Wang et al., 2007), and the expression patterns of *LEA* genes in response to various abiotic stresses have been investigated (Pedrosa et al., 2015; Wu et al., 2018). Currently, the functional mechanisms of LEA proteins in plant stress responses have become a popular research topic.

Tea plant [*Camellia sinensis* (L.) O. Kuntze] is an important economic crop which is widely cultivated worldwide (Mukhopadhyay et al., 2016). However, tea plant frequently experiences various abiotic stresses during its lifecycle, and cold and drought stresses are the main factors that significantly affect the yield and quality of tea products (Das et al., 2016; Wang et al., 2016). Recently, 2 LEA genes in tea plant were reported to be involved in the response to cold and drought stresses (Muoki et al., 2012; Paul et al., 2014), revealing the importance of *LEA* genes in tea plants under stress. However, the *LEA* gene family in tea plant has not been revealed, and the biological functions of the members of this family remain unknown, particularly in the tea plant's response to abiotic stresses. Here, *CsLEA* genes were identified from the tea plant genome via a genome-wide study, and a comprehensive analysis, which included individual analyses of sequence characteristics, phylogenetic relationships, gene structures, conserved motifs, protein interaction networks and transcription factor regulatory networks, was performed. In addition, expression profiles of *CsLEA* genes in different tissues of tea plant and in response to cold and dehydration stresses were measured. The results of this study should increase our understanding of the *CsLEA* gene family and provide a foundation for additional studies of the biological functions of the CsLEA proteins under abiotic stresses.

2. Materials and methods

2.1. Plant materials and stress treatments

Two-year-old cutting seedlings of tea plants [*C. sinensis* cv. ‘Longjingchangye’] were planted in an artificial climate chamber at Northwest A&F University (Yangling, China, N 34°16′, E 108°3′) under normal environmental conditions (12 h light/12 h dark photoperiod, 25 ± 2 °C ambient temperature and 70 ± 5% relative humidity). The tea plants were then exposed to cold and dehydration stress treatments, respectively. For cold treatment, tea plants were transferred to an artificial climate chamber maintained at 4 °C for 48 h. For dehydration treatment, the root of tea plants together with the matrix was immersed completely in the solution containing 20% (w/v) polyethylene glycol (PEG) 6000 for 48 h. And all stress treatments were completed under other growth conditions consistent. The first and second young leaves from about 120 tea plants were randomly collected at 0, 1, 2, 4, 8, 12, 24 and 48 h after stress treatments. Additionally, roots, stems, leaves and flowers from the tea plants grown in the artificial climate chamber under normal environmental conditions (mentioned above) were collected in October. For all conditions, three independent biological replicates were prepared. All the samples were immediately frozen in liquid nitrogen and stored at −80 °C for further analysis.

2.2. Total RNA isolation and cDNA reverse transcription

The total RNA from the sample was extracted using RNAiso Plus (TaKaRa, Dalian, China), and all genomic DNA contamination was eliminated with DNase I (TaKaRa, Dalian, China). The concentration and quality of the RNA were measured using a NanoDrop ND-1000 spectrophotometer (NanoDrop, Wilmington, USA). Afterward, 1 μg of total RNA was reverse transcribed to cDNA using a 5 × All-In-One RT MasterMix Kit (ABM, Richmond, Canada) in accordance with the manufacturer's protocol.

2.3. Identification of members of the CsLEA gene family in tea plant

The LEA protein sequences in Arabidopsis and rice were obtained from The Arabidopsis Information Resource (TAIR) (http://www.arabidopsis.org/) and the Rice Genome Annotation Project (RGAP) (http://rice.plantbiology.msu.edu/) databases and were used as queries for BLASTP searches against the tea plant genome (Wei et al., 2018; Xia et al., 2017). The obtained *CsLEA* genes were then rechecked and confirmed by querying the NCBI database (https://www.ncbi.nlm.nih.gov/).

2.4. Sequence analysis and phylogenetic tree construction

The physicochemical parameters, including molecular weight, theoretical isoelectric point (pI), instability index, aliphatic index and the grand average of hydropathicity (GRAVY), of the CsLEA proteins were analysed with the ExPASy ProtParam tool (https://web.expasy.org/protparam/). The WoLF PSORT online program (https://wolfpsort.hgc.jp/) was used to predict the subcellular localization of the CsLEA proteins. Multiple sequence alignment of the amino acid sequences of the LEA proteins in tea plant, Arabidopsis and rice was performed via of DNAMAN 6.0 with the default parameters. A phylogenetic tree was constructed using the neighbour-joining method with 1000 bootstrap replicates in the MEGA 6.0, and the diagram was drawn and retouched via EVOLVIEW (www.evolgenius.info/evolview).

2.5. Analysis of gene structure and protein conserved motifs

The exon-intron structures of the *CsLEA* genes were analysed by comparing coding sequences with their corresponding genomic sequences using Gene Structure Display Server 2.0 (http://gsds.cbi.pku.edu.cn). MEME 5.0.1 (http://meme-suite.org/tools/meme) was used to identify the conserved motifs of the CsLEA proteins with the following parameters: number of repetitions, any; maximum number of motifs, 10; and optimum motif widths, 6 to 50 amino acid residues.

2.6. Prediction of protein interaction networks and transcription factor regulatory networks

The functional interacting network models of CsLEA proteins were predicted using the web program STRING (http://string-db.org); the confidence parameter was set at a threshold of 0.40. The 600 bp sequences upstream from the start codon of the *CsLEA* genes were extracted from the tea plant genome (Wei et al., 2018; Xia et al., 2017); the sequence information is shown in Table S1. The transcription factors of the *CsLEA* genes then were predicted using the Plant Transcriptional Regulatory Map (PTRM, http://plantregmap.cbi.pku.edu.cn/regulation_prediction.php) with a threshold (for binding site prediction) p-value ≤ 1e-5, after which the transcription factor enrichment results were used to construct a transcription factor regulatory network via Cytoscape 3.6 software.

2.7. Quantitative RT-PCR (qRT-PCR) analysis of CsLEA genes

An expression analysis of the *CsLEA* genes was performed via qRT-PCR using SYBR® *Premix Ex Taq*™ II (TaKaRa, Dalian, China) on a Bio-Rad IQ5 Real-Time PCR System (Bio-Rad, Hercules, USA) in accordance with the manufacturer's instructions. The *CsPTB* gene (GeneBank: GAAC01052498.1) was used as an internal control (Hao et al., 2014); all primers used for qRT-PCR analysis are listed in Table S2. The experiments were repeated in triplicate, and the relative gene expression levels were calculated based on the threshold cycle using the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001). The HeatMaps of the gene expression were visualized using Cluster 3.0 and TreeView software.

Table 1
Summary information of the physiological and biochemical properties of CsLEA proteins.

Gene name	Gene ID[a]	Animo acids	MW (kDa)	pI	GRAVY	Instability index	Aliphatic index	Subcellular localization [b]
CsLEA1	CSA003726.1	96	10.44	9.99	−0.42	45.54	76.25	chlo: 7, cyto: 7
CsLEA2	TEA013594.1	124	13.54	5.92	−0.47	39.22	75.48	chlo: 10, mito: 4
CsLEA3	TEA Scaffold426:839639:840621:	175	19.80	9.37	0.02	21.71	110.80	cyto: 6, nucl: 2.5, pero: 2, cysk_nucl: 2, chlo: 1, E.R.: 1, golg: 1
CsLEA4	TEA006416.1/CSA006938.1	96	10.35	9.99	−0.53	37.17	71.25	mito: 6, chlo: 4, cyto: 3, nucl: 1
CsLEA5	CSA009009.1	106	11.39	7.11	−1.05	51.09	45.38	mito: 6, nucl: 3, cyto: 3, chlo: 2
CsLEA6	TEA011337.1/CSA030792.1	318	35.36	4.86	−0.37	19.90	96.23	cysk: 7, cyto: 4, chlo: 1, nucl: 1, golg: 1
CsLEA7	TEA014379.1/CSA031822.1	153	16.64	4.93	−0.01	12.22	103.86	cyto: 7, chlo: 3, nucl: 2, mito: 1, plas: 1
CsLEA8	TEA031400.1/CSA027557.1	250	25.90	4.64	−0.23	31.65	86.32	cyto: 9, chlo: 3, mito: 1, cysk_nucl: 1
CsLEA9	CSA000643.1	156	16.46	5.43	−1.23	36.16	32.88	mito: 8, nucl: 4, cyto: 1, plas: 1
CsLEA10	TEA032725.1	201	21.05	7.97	−1.32	14.17	23.28	nucl: 11, cyto: 1, plas: 1, extr: 1
CsLEA11	TEA010666.1/CSA004403.1	252	28.71	5.33	−1.47	59.23	47.58	nucl: 13, cysk: 1
CsLEA12	CSA019232.1	72	7.55	5.62	−0.10	26.91	82.64	cyto: 11, chlo: 3
CsLEA13	CSA012431.1	86	9.83	9.69	−0.39	35.85	85.00	chlo: 11, mito: 2, cyto: 1
CsLEA14	TEA026687.1/CSA030468.1	406	43.37	5.15	−0.86	26.62	39.56	nucl: 8.5, nucl_plas: 5, chlo: 2, mito: 2, cyto: 1
CsLEA15	TEA032830.1/CSA034141.1	221	24.55	10.16	0.01	25.15	97.87	cyto: 9, E.R.: 2, chlo: 1, extr: 1, vacu: 1
CsLEA16	TEA003016.1/CSA010151.1	187	20.69	7.64	−0.99	40.88	58.61	pero: 10, nucl: 3, cyto: 1
CsLEA17	TEA028150.1	128	14.08	8.95	−0.99	33.19	55.86	mito: 6, cyto: 4, chlo: 2, nucl: 2
CsLEA18	TEA011974.1	154	16.19	4.83	−1.21	31.56	29.61	nucl: 7, mito: 6, chlo: 1
CsLEA19	CSA033014.1	85	9.12	5.86	−1.02	46.98	43.88	mito: 9, nucl: 4, cyto: 1
CsLEA20	TEA032971.1/CSA010673.1	225	24.77	9.51	0.01	24.80	99.16	cyto: 10, E.R.: 3, vacu: 1
CsLEA21	TEA028236.1/CSA026452.1	154	15.76	6.51	−0.77	24.54	39.55	nucl: 7, mito: 5, chlo: 1, cyto: 1
CsLEA22	TEA033745.1	255	27.92	4.98	−1.06	25.46	59.06	cyto: 3, extr: 3, E.R.: 2.5, E.R._plas: 2.5, mito: 2, plas: 1.5, chlo: 1, nucl: 1
CsLEA23	TEA000416.1/CSA020543.1	378	40.79	5.42	−0.71	26.93	49.26	extr: 6, vacu: 4, mito: 2, nucl: 1, golg: 1
CsLEA24	TEA005221.1/CSA031121.1	333	36.35	5.04	−1.11	46.42	51.44	chlo: 8, nucl: 3, mito: 2, cyto: 1
CsLEA25	TEA030980.1/CSA009868.1	210	23.57	9.92	−0.23	37.53	84.95	cyto: 4.5, golg: 3, cyto_pero: 3, nucl: 2, vacu: 2, mito: 1, E.R.: 1
CsLEA26	CSA035676.1	78	8.59	5.45	−1.52	44.99	35.00	nucl: 6, cyto_nucl: 5, mito: 4, cyto: 2, extr: 2
CsLEA27	TEA018323.1/CSA002944.1	206	22.68	9.71	0.20	33.83	110.10	cyto: 8, vacu: 2, E.R.: 2, chlo: 1, plas: 1
CsLEA28	TEA021072.1/CSA036017.1	194	21.30	10.02	−0.04	33.60	85.88	cyto: 7, nucl: 2, E.R.: 2, chlo: 1, mito: 1, extr: 1
CsLEA29	TEA007147.1/CSA021249.1	211	23.19	9.66	0.07	58.19	86.26	chlo: 7, mito: 2, vacu: 2, nucl: 1, plas: 1, extr: 1
CsLEA30	TEA005685.1/CSA003226.1	197	21.50	9.37	0.29	33.85	88.58	chlo: 6, mito: 4, extr: 1, vacu: 1, E.R.: 1, cyto_nucl: 1
CsLEA31	TEA015290.1	213	24.02	9.82	−0.34	47.37	78.73	cyto: 4, vacu: 3, nucl: 2, E.R.: 2, golg: 2, plas: 1
CsLEA32	TEA015292.1	210	23.37	9.95	−0.11	33.95	87.71	cyto: 5, vacu: 3, E.R.: 2, golg: 2, nucl: 1, mito: 1
CsLEA33	TEA015293.1	212	23.59	9.64	−0.03	48.73	86.46	cyto: 5, golg: 3, vacu: 2, plas: 1.5, E.R._plas: 1.5, nucl: 1, mito: 1

[a] Gene IDs are from *C. sinensis* var. *sinensis* genome (marked with TEA) and *C. sinensis* var. *assamica* genome (marked with CSA).

[b] chlo: chloroplast; cyto: cytoplasm; ER: endoplasmic reticulum; extr: extracellular; golg: Golgi apparatus; mito: mitochondrion; nucl: nucleus; pero: peroxide; plas: plasma membrane; vacu: vacuole.

3. Results

3.1. Identification and characterization of the CsLEA gene family

In total, 33 *CsLEA* genes, named *CsLEA1* to *CsLEA33*, were identified from the genome of tea plant (Table S3). The results of the physiological and biochemical properties analysis showed that the lengths of the CsLEA proteins ranged from 72 to 406 amino acids, the molecular weights were between 7.55 kDa and 43.37 kDa, and the pI values ranged from 4.64 to 10.16. With the exception of CsLEA3, 15, 20, 27, 29 and 30, most CsLEA proteins presented GRAVY values of less than 0, implying that they all are hydrophilic proteins. More detailed information, including their instability index, aliphatic index and subcellular localization, is listed in Table 1.

3.2. Phylogenetic analysis of CsLEA proteins

To investigate the phylogenetic relationship of LEA proteins among tea plant, Arabidopsis and rice, a phylogenetic tree was constructed using the neighbour-joining method. As shown in Fig. 1, the LEA proteins were classified into eight groups, including LEA_1, LEA_2, LEA_3, LEA_4, LEA_5, LEA_6, dehydrin and SMP groups, while the 33 CsLEA proteins appeared in only seven groups. There were no members in the LEA_6 group. In addition, the LEA_2 and LEA_4 groups were the largest, with thirteen and eight LEA members in tea plant, respectively. By contrast, the LEA_5 group had only one CsLEA member, CsLEA26.

Twenty-three other species (Li, 2016) in which members of the LEA family have been identified were comparatively analysed to further assess the family composition of CsLEA proteins, and the results showed that the distribution of the CsLEA members was similar in proportion to that of other higher plants, especially angiosperms. Specifically, the LEA_2 and LEA_4 group had significantly more members than did the other groups, although this trend was not completely manifested in all higher plants. By contrast, the structure of the LEA members in lower plants and animals is relatively simple. Moreover, no LEA_6 members were identified in Solanaceae plants, which, in addition to tea plant, are Asterids; this finding suggested that gene loss events occurred during the process of Asterid plant evolution (Fig. 2).

3.3. Gene structures and conserved motifs within the CsLEA family

To further analyse the evolutionary relationships of the *CsLEA* gene family, the exon-intron organization of the *CsLEA* genes was investigated (Fig. 3). The LEA_5 group had no introns; the LEA_1, LEA_3 and dehydrin groups had only 1 intron; and the other groups had 0-2 introns. Generally, the *CsLEA* genes in the same group had similar exon-intron feature, which supports their close phylogenetic relationship and group classification. In addition, conserved motifs were predicted and analysed using the MEME tool to understand the structural features of the CsLEA proteins. Ten motifs describing the details of the CsLEA

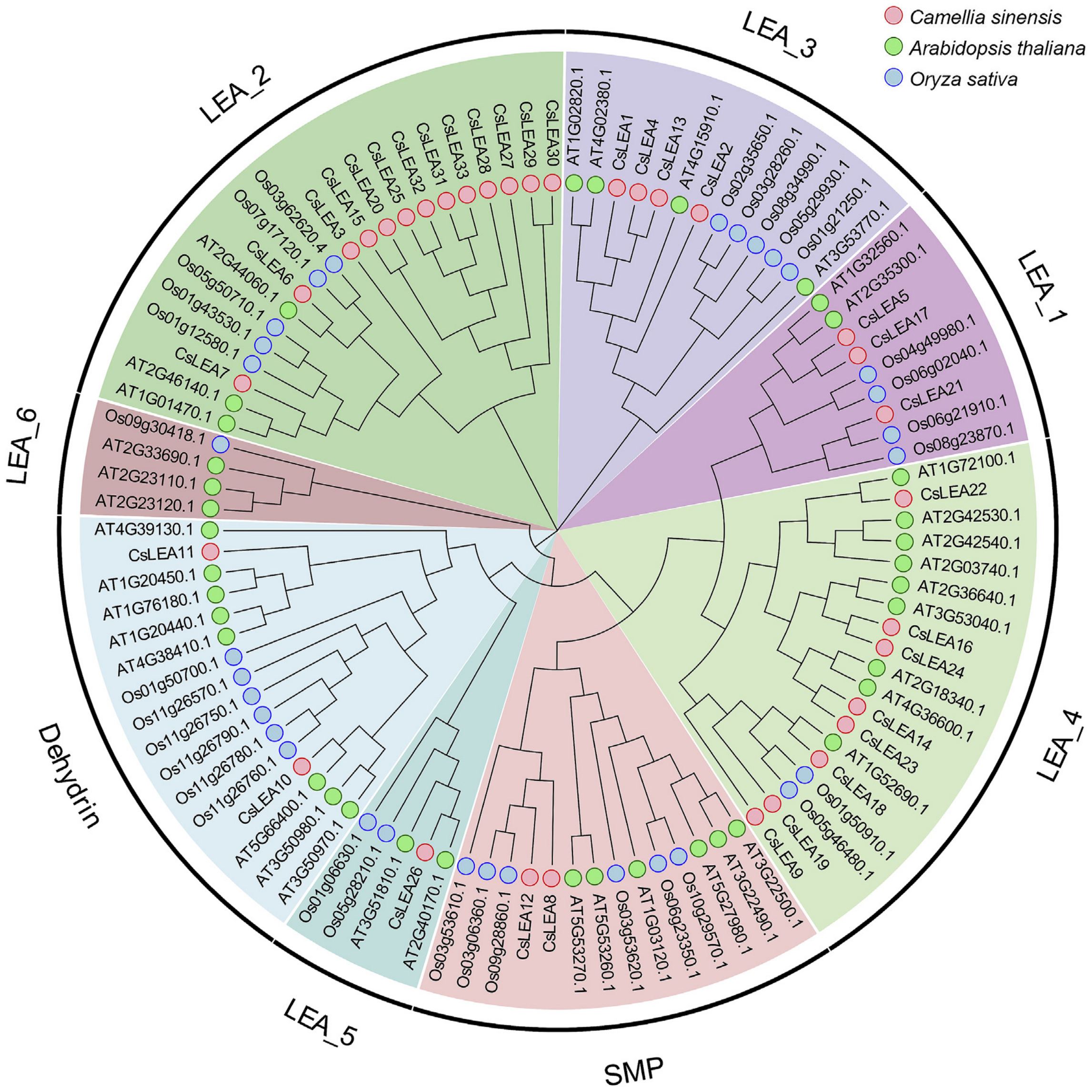


Fig. 1. Phylogenetic tree of LEA proteins in tea plant, Arabidopsis and rice constructed using the neighbour-joining method.

proteins were ultimately predicted (Fig. 4 and Figure S1); the regular expression levels of the conserved motifs are listed in Table S4. The 33 CsLEA proteins did not share high similarity, but the LEA proteins within the same group had a similar motif organization, which further supports the results of the phylogenetic analysis. On the other hand, significant differences in motifs among the different groups of CsLEA proteins hint at protein functional diversity.

3.4. Protein interaction network of CsLEA proteins

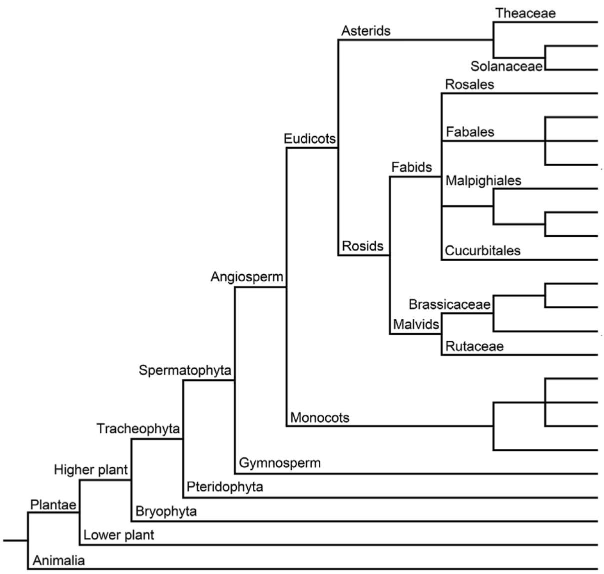
In this study, the CsLEA proteins were investigated in an Arabidopsis association model using STRING software to identify the functional and physical interactions. The results in Fig. 5 show that 32 CsLEA proteins were mapped to 18 known Arabidopsis proteins that participate in the interaction network; CsLEA22 was the lone exception. Most CsLEA proteins were clearly associated with each other and are involved in the CBF/DREB signalling pathway by interacting with CBF2, ICE1, DREB1A, ABI1 or ABF4. In contrast, some CsLEA proteins were independent, such as CsLEA14, 15, 20, 23, 25, 27, 28, 29, 30, 31, 32 and 33.

3.5. Transcription factor regulatory network of CsLEA genes

To identify the potential transcriptional regulatory network of the *CsLEA* gene family, 600 bp upstream sequences from the 33 *CsLEA* genes were extracted and analysed, and the results showed that 108 transcription factors that participated in the regulation of 31 *CsLEA* genes, including numerous ERF, MYB, bZIP and bHLH transcription factor family members and a few other family members, were enriched (Fig. 6 and Table S5). Among these transcription factors, the ERF family members were most abundant, and most, including CBF1, CBF2, CBF3 and CBF4, belonged to the DREB subfamily. In addition, ABF1, ABF2, ABF3 and ABI5 of the bZIP family were also identified in the regulatory network, which is consistent with the protein interaction network of the CsLEA proteins (Fig. 5). Interestingly, three BPC family transcription factors, BPC1, BPC5 and BPC6, were shown to universally regulate the *CsLEA* genes, especially members of the LEA_4 group.

3.6. Expression analysis of CsLEA genes in different tissues of tea plant

The expression patterns of *CsLEA* genes in roots, stems, leaves and flowers were investigated by RT-PCR. As shown in Fig. 7 and Table S6, the *CsLEA* genes were differentially expressed in different tissues of tea



Species	Group								Other	Total
	LEA_1	LEA_2	LEA_3	LEA_4	LEA_5	LEA_6	Dehydrin	SMP		
Camellia sinensis	3	13	4	8	1	-	2	2	-	33
Solanum tuberosum	5	5	3	6	2	-	5	2	1	29
Solanum lycopersicum	5	5	3	3	2	-	6	3	-	27
Prunus mume	3	6	3	5	3	-	6	3	1	30
Medicago truncatula	4	2	2	5	1	2	3	4	-	23
Glycine max	2	6	8	6	3	2	4	5	-	36
Arachis hypogaea	2	1	2	7	3	1	1	2	-	19
Ricinus communis	2	2	3	7	2	2	5	4	-	27
Populus tomentosa	4	61	4	6	1	2	7	2	-	87
Populus trichocarpa	3	4	5	26	1	2	10	2	-	53
Cucumis sativus	2	-	4	-	2	2	4	3	62	79
Brassica rapa	5	5	13	2	3	2	11	8	4	53
Brassica napus	9	10	13	25	7	5	23	16	-	108
Arabidopsis thaliana	3	3	4	18	2	3	10	6	2	51
Citrus sinensis	4	41	5	7	3	-	6	4	2	72
Oryza sativa	4	7	5	8	2	1	6	6	-	39
Brachypodium distachyon	5	5	8	1	3	1	9	6	2	40
Phyllostachys edulis	1	5	5	5	-	1	6	-	-	23
Zea mays	4	6	7	3	2	1	5	4	-	32
Pinus tabuliformis	2	1	6	3	1	-	6	4	-	23
Selaginella moellendorffii	1	3	4	-	3	-	-	4	-	15
Physcomitrella patens	-	3	1	2	3	-	3	2	4	18
Chlamydomonas reinhardtii	-	-	-	-	1	-	-	-	-	1
Polypedilum vanderplanki	-	-	-	27	-	-	-	-	-	27

Fig. 2. Summary of the LEA gene family among 24 species.

plant and were grouped into 2 classes. Among these *CsLEA* genes in Class I, 8 were highly expressed in the roots, 5 and 9 were highly expressed in the stems and leaves, respectively; except for *CsLEA3*, all others were expressed in low amounts in the flowers. In contrast, 11 *CsLEA* genes in Class II, mainly members of the LEA_4 and SMP groups, were expressed more in the flowers than in other organs, although some of these genes were also slightly expressed in other tissues.

3.7. Expression analysis of CsLEA genes under cold and dehydration stress

Under cold stress, the expression of 20 *CsLEA* genes was significantly upregulated, but the expression of a subset of those genes briefly decreased at 1, 2 and 48 h; in contrast, the expression of the remaining 13 *CsLEA* genes was significantly downregulated, especially that of the members of the LEA_4 group (Fig. 8A and Table S7). Interestingly, similar expression trends were also observed under dehydration stress; *CsLEA20*, *CsLEA26* and all LEA_4 group genes were

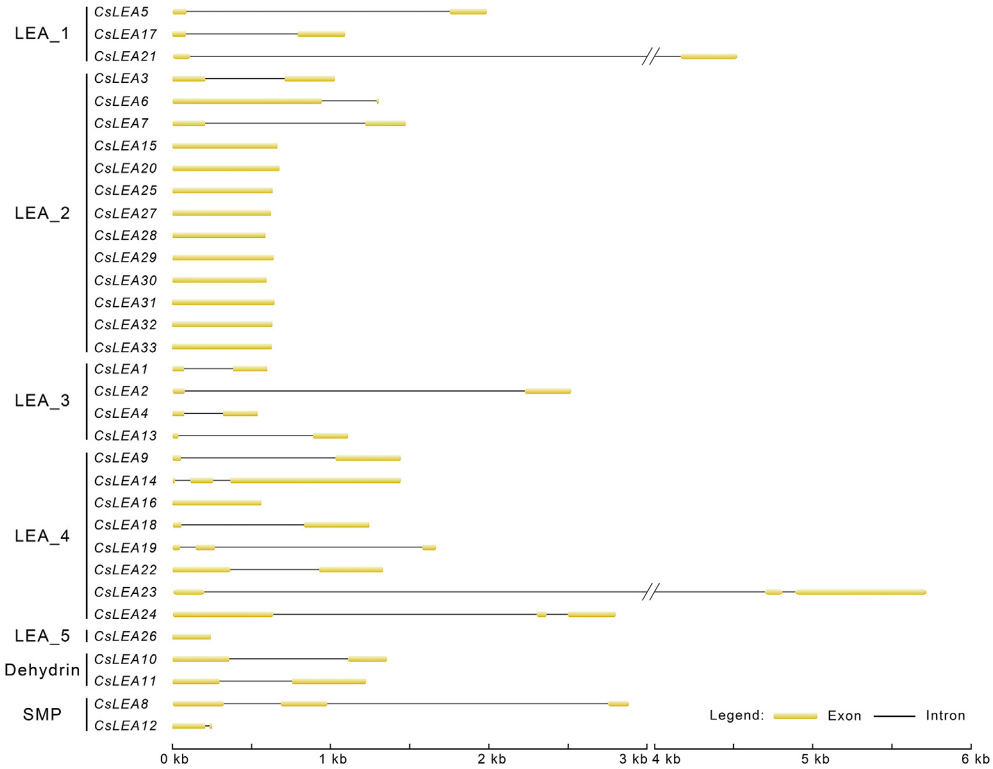


Fig. 3. Exon-intron structure analyses of members of the *CsLEA* gene family. The yellow sections represent exons, and the grey parts indicate introns. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

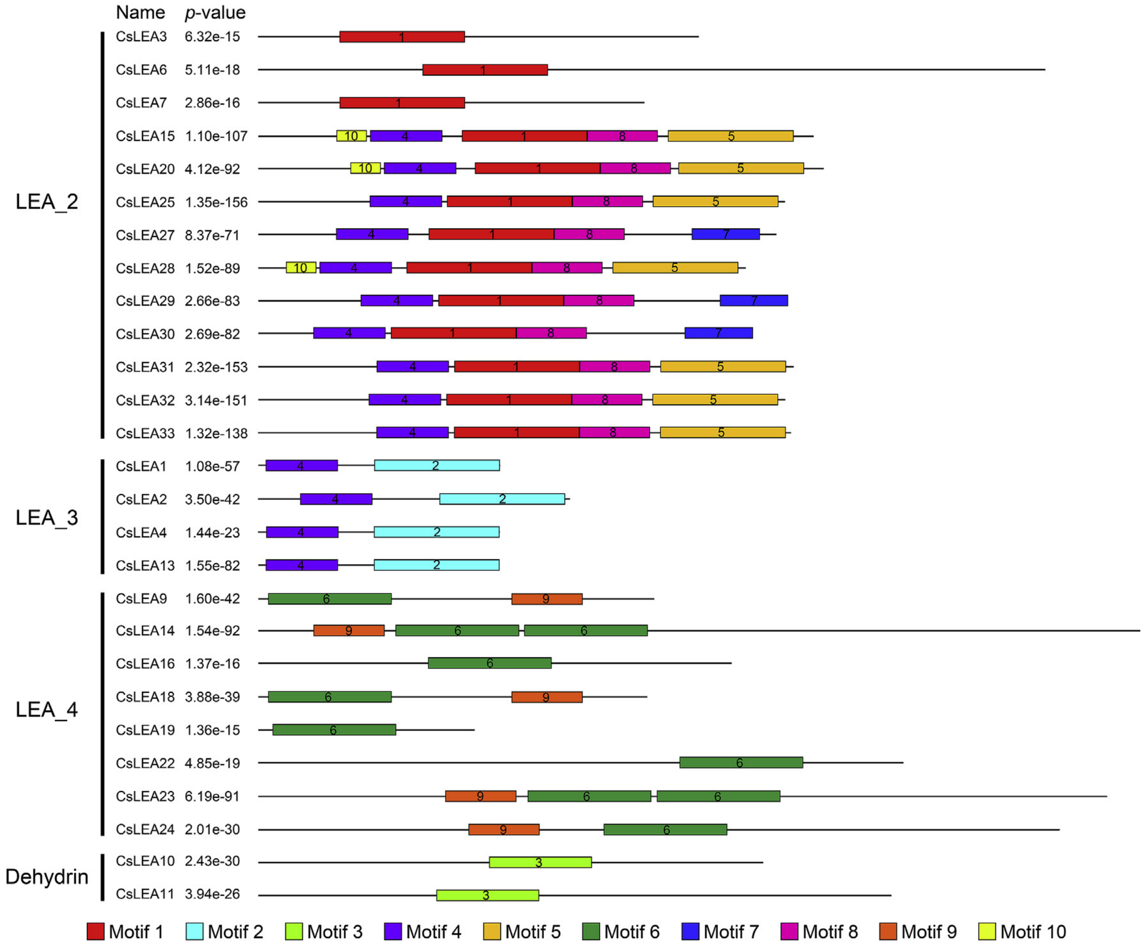


Fig. 4. Distribution of conserved motifs within CsLEA proteins. Putative motifs are represented by a number in a coloured box.

significantly downregulated; and the expression of the other *CsLEA* genes was significantly upregulated (Fig. 8B and Table S8). In addition, the expression trends of the same group of *CsLEA* genes were essentially consistent under cold and dehydration stresses. These results reveal that almost all *CsLEA* genes are involved in the response of tea plant to cold and dehydration stresses, and this response may depend on similar mechanisms.

4. Discussion

LEA proteins are typically associated with protective functions that play crucial roles during embryonic development and response to abiotic stresses in plants (Shih et al., 2008). To date, members of *LEA* gene families in many plants have been reported, but they have not been revealed in tea plant. Excitingly, publication of the tea plant genome rendered the identification and characterization of the *CsLEA* gene family possible (Wei et al., 2018; Xia et al., 2017). In the present study, 33 *CsLEA* genes were identified in tea plant, which is consistent with the number of *LEA* members in tomato (Cao and Li, 2015), potato (Charfeddine et al., 2015), Chinese plum (Du et al., 2013), and maize (Li and Cao, 2016), etc., but lower than that in Arabidopsis (Hundertmark and Hincha, 2008), poplar (Lan et al., 2013), edible rape (Liang et al., 2016), sweet orange (Pedrosa et al., 2015), etc. Previous phylogenetic analyses classified plant LEA proteins into at least eight groups, including LEA_1, LEA_2, LEA_3, LEA_4, LEA_5, LEA_6, dehydrin and SMP groups. However, the CsLEA proteins were classified into only seven groups, among which the LEA_6 group was not identified in tea plant. Interestingly, the lack of LEA_6 group members has also been reported in other Asterid plants such as tomato and potato (Cao and Li, 2015; Charfeddine et al., 2015), suggesting that gene loss events occurred during the process of Asterid plant evolution; however, this supposition requires more genetic evolution information for confirmation. In addition, studies have shown that the greatest number of LEA_2 group genes occurs in higher plants (Pedrosa et al., 2015), which was also confirmed by our results, as the LEA_2 group had significantly more members than did the other groups in tea plant. However, advantages of the quantity of LEA_2 members have not been described in previous genome investigations of Arabidopsis (Hundertmark and Hincha, 2008), rice (Wang et al., 2007) or poplar (Lan et al., 2013). These inconsistent results may be explained in part by the improvement of higher plant genome annotations and the duplication of atypical LEA proteins in the LEA_2 group (Pedrosa et al., 2015). These findings suggest that the LEA protein family in higher plants may be larger and much more complex than that currently described, especially the LEA_2 group proteins.

Accumulating evidence has demonstrated that genes related to the stress response usually contain fewer introns (Jeffares et al., 2008). Our results also showed that all *CsLEA* genes have no more than two introns, which is similar to the exon-intron organization of LEA genes in other plants such as Chinese plum (Du et al., 2013), poplar (Lan et al., 2013), tomato (Cao and Li, 2015) and potato (Charfeddine et al., 2015). A low abundance of introns contributes to the transcriptional regulation of *LEA* genes under stress conditions, which possibly explains why *CsLEA* genes are rapidly and strongly induced or inhibited under cold and dehydration stresses. In addition, we found that *CsLEA* genes in the same group have similar exon-intron features, which supports their close phylogenetic relationship and group classification. Similar results concerning the conserved motif composition of CsLEA proteins have

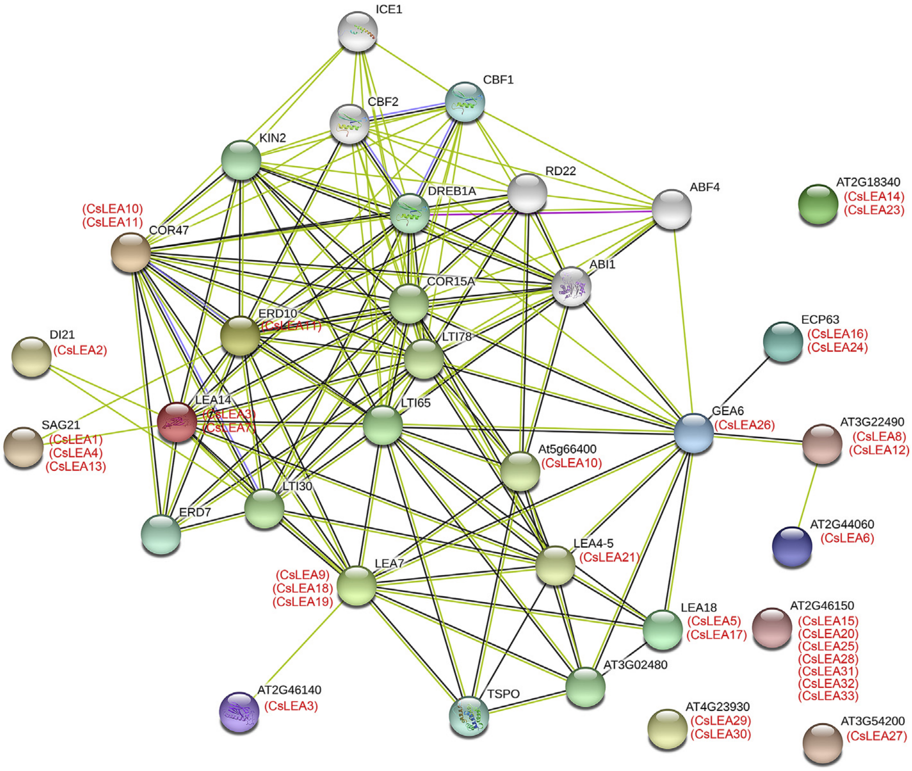


Fig. 5. Putative protein interaction network of CsLEA proteins in tea plant. Homologous proteins in tea plant and Arabidopsis are shown in red and black, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

also been reported, which further supports the results of phylogenetic analyses. On the other hand, many studies have indicated that LEA proteins in different groups exhibit low similarity (Shih et al., 2008). Our results also showed that the CsLEA proteins in each group have distinct conserved motifs that have also been identified in other plant species, such as pine (Gao and Lan, 2016), *Gastrodia elata* (Zeng et al., 2018), tomato (Cao and Li, 2015) and potato (Charfeddine et al., 2015), implying that CsLEA proteins are functional LEA proteins that have group-specific functions; furthermore, members of the same group might originate from gene duplication within that group, while the different groups may have evolved from different ancestors (Altunoglu et al., 2016; Ling et al., 2016; Pedrosa et al., 2015; Zeng et al., 2018).

Recent studies have shown significant histological specificity in the transcription of plant *LEA* genes. For example, most of the *LEA* genes in maize show distinct tissue-specific expression in 60 tissues (Li and Cao, 2016), and 9 potato *LEA* genes are highly expressed in the vegetative organs (Charfeddine et al., 2015). In cassava, the expression levels of 7 *MeLEA* genes were not detected or were at low levels in different

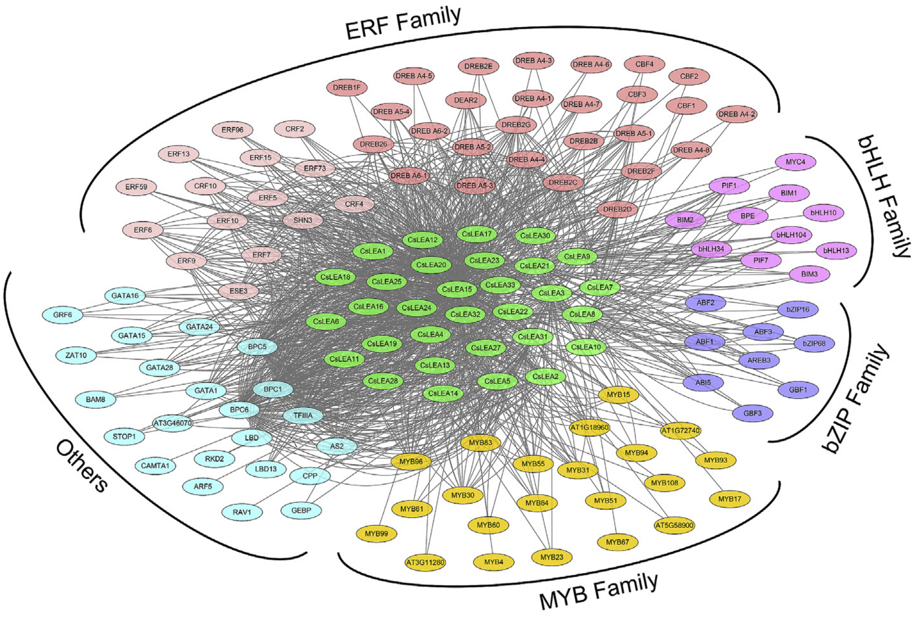


Fig. 6. Transcription factor regulatory network of *CsLEA* genes. The 600 bp upstream sequence from each *CsLEA* gene was extracted and used to construct the transcriptional regulatory network via the PTRM tool and Cytoscape 3.6 software.

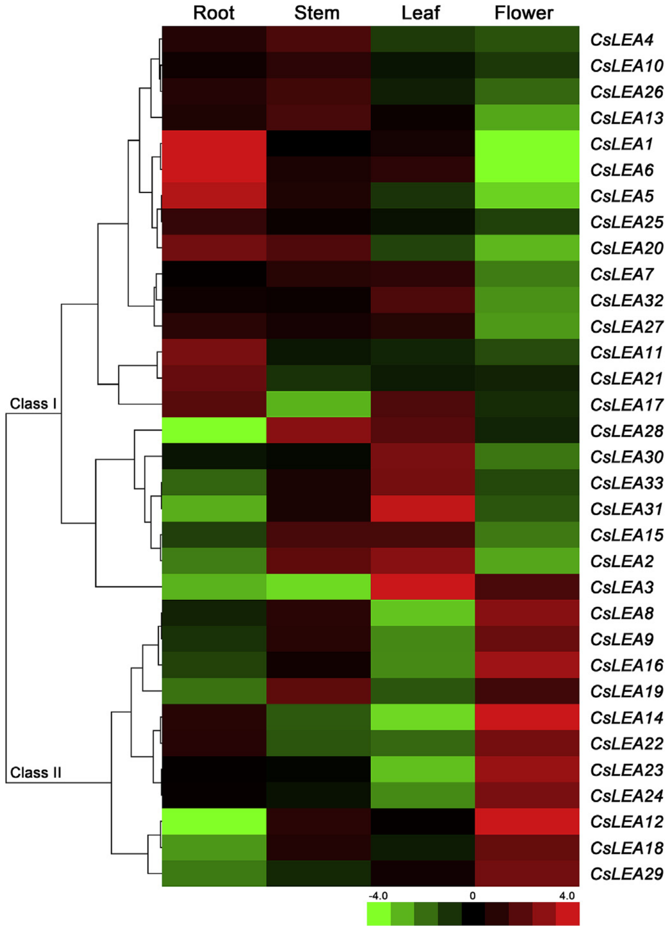


Fig. 7. Expression patterns of *CsLEA* genes in different tissues of tea plant. The HeatMap was generated by Cluster 3.0 and TreeView software using the *CsLEA* gene expression data, and normalized $\log_2$-transformed values were used with hierarchical clustering. The red and green colours indicate high or low transcript abundances, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

tissues, and 8, 7 and 5 *MeLEA* genes were highly expressed in storage roots, stems and leaves, respectively (Wu et al., 2018). In the present study, we also found that the *CsLEA* genes were specifically expressed in one or more tissues, which is similar to the results from recent reports in tomato (Cao and Li, 2015), edible rape (Liang et al., 2016) and sweet orange (Pedrosa et al., 2015), suggesting that the *CsLEA* genes are extensively involved in the growth and development of tea plant and may develop different biological functions in different tissues. It is widely believed that LEA proteins constitute an important response to different abiotic stresses and are usually induced under stress conditions such as cold, drought and high salinity (Huang et al., 2016). As expected, the expression of many *CsLEA* genes were significantly induced under cold and dehydration stresses, and their trends were essentially consistent, which is similar to the results of previous investigations in other plants (Li and Cao, 2016; Wu et al., 2018; Zeng et al., 2018); these findings imply that these CsLEA proteins are critical in the tea plant response to cold and dehydration stresses. Interestingly, the expression levels of the *CsLEA* genes in the LEA_4 group were significantly downregulated under cold and dehydration stresses, and these genes were specifically expressed in the tea plant flowers, suggesting that LEA_4 group CsLEA proteins are critical for reproductive growth but not for stress responses, although this needs further confirmation.

Previous investigations have confirmed that LEA proteins, as downstream factors of the CBF/DREB signalling pathway, are involved in plant responses to cold and drought stresses (Kaur and Asthir, 2017; Kosova et al., 2007). For example, DREB1A/CBF3 transcription factors can simultaneously activate the expression of many *LEA* genes, resulting in multiple improvements in Arabidopsis stress tolerance (Liu et al., 2000). Overexpression of two DREB/CBF transcription factors improved the stress tolerance of transgenic wheat by increasing the expression of many LEA genes (Morran et al., 2011). Our protein interaction prediction results show that many CsLEA proteins are involved in the CBF/DREB signalling pathway, implying that the expression of *CsLEA* genes is regulated by DREB/CBF transcription factors under stress. Interestingly, the greatest quantify of transcription factors identified in the transcriptional regulatory network of the *CsLEA* gene family were the DREB/CBF type, which further validates our above speculation. In addition, increasing amounts of evidence have indicated that the expression of LEA genes is also regulated by transcription factors of the MYB and bZIP families. For instance, cotton GbMYB5 activates the expression of LEA genes to maintain the stabilization of plant cells under drought stress (Chen et al., 2015), and Arabidopsis ABI5 regulates the accumulation of LEA proteins during dark-induced leaf senescence (Su et al., 2016). We also found that many MYB, bZIP and bHLH family transcription factors are involved in the regulation of *CsLEA* genes, which is similar to the results of analyses of *cis*-acting elements of promoters in tomato (Cao and Li, 2015), maize (Li and Cao, 2016) and sweet orange (Pedrosa et al., 2015). Furthermore, the *CsLEA* genes of the LEA_4 group are generally regulated by multiple BPC transcription factors that are critical for plant growth and development, especially reproductive growth (Monfared et al., 2011), which also accurately explains the results of LEA_4 group *CsLEA* gene-specific expression in the tea plant flowers. Overall, these results suggest that a potential transcriptional regulatory network, including DREB/CBF, MYB, bZIP, bHLH, BPC and other transcription factors, is directly associated with the expression of *CsLEA* genes, which may be ubiquitous and play important roles in tea plant growth, development and stress responses.

In summary, this study is the first to comprehensively and systematically analyse members of the LEA gene family in tea plant. In total, 33 *CsLEA* genes were identified in the tea plant genome, and bioinformatic and expression profile analyses were performed to determine their potential functions. The results indicate that *CsLEA* genes are actively involved in regulating tea plant growth, development and responses to cold and dehydration stress, and these processes are closely related to a transcriptional regulatory network involving DREB/CBF, MYB, bZIP, bHLH, BPC and other transcription factors. These results will provide a solid foundation for additional functional studies investigating CsLEA proteins and contribute to an improved understanding of the mechanisms of stress tolerance mediated by CsLEA proteins in tea plant.

Conflicts of interest

The authors declare no conflict of interest.

CRediT authorship contribution statement

Weidong Wang: Conceptualization, Data curation, Formal analysis, Writing – original draft. **Tong Gao:** Data curation, Formal analysis, Writing – original draft. **Jiangfei Chen:** Data curation, Formal analysis. **Jiankun Yang:** Formal analysis. **Huiyu Huang:** Formal analysis. **Youben Yu:** Conceptualization, Writing – review & editing.

Acknowledgements

This research was funded by the China Postdoctoral Science Foundation (2016M602873), the Earmarked Fund for Modern Agro-industry Technology Research System (CARS-19), and the special fund for University-Supported Extension Model (XTG2018-04).

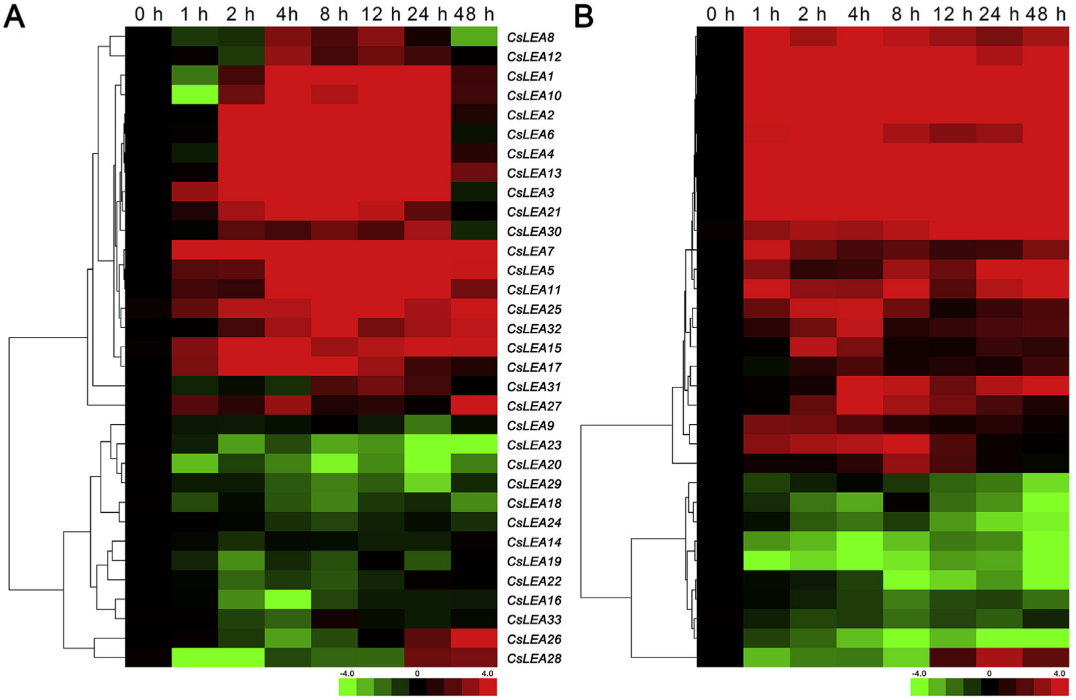


Fig. 8. Expression patterns of *CsLEA* genes in tea plant under cold (A) and dehydration (B) stresses. The HeatMap was generated by Cluster 3.0 and TreeView software using the *CsLEA* gene expression data, and normalized log$_2$-transformed values were used with hierarchical clustering. The red and green colours indicate higher or lower transcript abundances compared with those at 0 h, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.plaphy.2018.12.009.

References

Altunoglu, Y.C., Baloglu, P., Yer, E.N., Pekol, S., Baloglu, M.C., 2016. Identification and expression analysis of *LEA* gene family members in cucumber genome. Plant Growth Regul. 80, 225–241.

Browne, J., Tunnacliffe, A., Burnell, A., 2002. Anhydrobiosis - plant desiccation gene found in a nematode. Nature 416, 38.

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

Charfeddine, S., Saidi, M.N., Charfeddine, M., Gargouri-Bouzid, R., 2015. Genome-wide identification and expression profiling of the late embryogenesis abundant genes in potato with emphasis on dehydrins. Mol. Biol. Rep. 42, 1163–1174.

Chen, T., Li, W., Hu, X., Guo, J., Liu, A., Zhang, B., 2015. A cotton MYB transcription factor, GbMYB5, is positively involved in plant adaptive response to drought stress. Plant Cell Physiol. 56, 917–929.

Close, T.J., Lammers, P.J., 1993. An osmotic stress protein of cyanobacteria is immunologically related to plant dehydrins. Plant Physiol. 101, 773–779.

Das, A., Mukhopadhyay, M., Mondal, T.K., 2016. Generation and characterization of expressed sequence tags in young roots of tea (*Camellia assamica*). Biol. Plant. 60, 48–54.

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*. Mol. Biol. Rep. 40, 1937–1946.

Dure, L.R., Greenway, S.C., Galau, G.A., 1981. Developmental biochemistry of cottonseed embryogenesis and germination: changing messenger ribonucleic acid populations as shown by in vitro and in vivo protein synthesis. Biochemistry 20, 4162–4168.

Gao, J., Lan, T., 2016. Functional characterization of the late embryogenesis abundant (LEA) protein gene family from *Pinus tabuliformis* (Pinaceae) in *Escherichia coli*. Sci. Rep. 6, 19467.

Hand, S.C., Menze, M.A., Toner, M., Boswell, L., Moore, D., 2011. LEA proteins during water stress: not just for plants anymore. *Annu.* Rev. Physiol. 73, 115–134.

Hao, X., Horvath, D.P., Chao, W.S., Yang, Y., Wang, X., Xiao, B., 2014. Identification and evaluation of reliable reference genes for quantitative real-time PCR analysis in tea plant (*Camellia sinensis* (L.) O. Kuntze). Int. J. Mol. Sci. 15, 22155–22172.

Hatanaka, R., Gusev, O., Cornette, R., Shimura, S., Kikuta, S., Okada, J., Okuda, T., Kikawada, T., 2015. Diversity of the expression profiles of late embryogenesis abundant (LEA) protein encoding genes in the anhydrobiotic midge *Polypedilum vanderplanki*. Planta 242, 451–459.

Hincha, D.K., Thalhammer, A., 2012. LEA proteins: IDPs with versatile functions in cellular dehydration tolerance. Biochem. Soc. Trans. 40, 1000–1003.

Huang, Z., Zhong, X.J., He, J., Jin, S.H., Guo, H.D., Yu, X.F., Zhou, Y.J., Li, X., Ma, M.D., Chen, Q.B., Long, H., 2016. Genome-Wide identification, characterization, and Stress-Responsive expression profiling of genes encoding LEA (Late embryogenesis abundant) proteins in moso bamboo (*Phyllostachys edulis*). PLoS One 11, e165953.

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., Bahler, J., 2008. Rapidly regulated genes are intron poor. Trends Genet. 24, 375–378.

Kaur, G., Asthir, B., 2017. Molecular responses to drought stress in plants. Biol. Plant. 61, 201–209.

Kosova, K., Vitamvas, P., Prasil, I.T., 2007. The role of dehydrins in plant response to cold. Biol. Plant. 51, 601–617.

Kosova, K., Vitamvas, P., Prasil, I.T., 2014. Wheat and barley dehydrins under cold, drought, and salinity - what can LEA-II proteins tell us about plant stress response? Front. Plant Sci. 5, 343.

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

Li, X., 2016. Molecular Evolutionary Research of the Plant *LEA* Gene Family and Preliminary Functional Analysis of an Atypical *LEA* Gene in Arabidopsis. Jangsu University.

Li, X., Cao, J., 2016. Late embryogenesis abundant (*LEA*) gene family in maize: identification, evolution, and expression profiles. Plant Mol. Biol. Rep. 34, 15–28.

Liang, Y., Xiong, Z., Zheng, J., Xu, D., Zhu, Z., Xiang, J., Gan, J., Raboanatahiry, N., Yin, Y., Li, M., 2016. Genome-wide identification, structural analysis and new insights into late embryogenesis abundant (*LEA*) gene family formation pattern in *Brassica napus*. Sci. Rep. 6, 24265.

Ling, H., Zeng, X., Guo, S., 2016. Functional insights into the late embryogenesis abundant (LEA) protein family from *Dendrobium officinale* (Orchidaceae) using an *Escherichia coli* system. Sci. Rep. 6, 39693.

Liu, H., Yu, C., Li, H., Ouyang, B., Wang, T., Zhang, J., Wang, X., Ye, Z., 2015a. Overexpression of *ShDHN*, a dehydrin gene from *Solanum habrochaites* enhances tolerance to multiple abiotic stresses in tomato. Plant Sci. 231, 198–211.

Liu, Q., Zhao, N.M., Yamaguch-Shinozaki, K., Shinozaki, K., 2000. Regulatory role of DREB transcription factors in plant drought, salt and cold tolerance. Chin. Sci. Bull. 45, 970–975.

Liu, Y., Liang, J., Sun, L., Yang, X., Li, D., 2016. Group 3 LEA protein, ZmLEA3, is involved in protection from low temperature stress. Front. Plant Sci. 7, 1011.

Liu, Y., Xie, L., Liang, X., Zhang, S., 2015b. *CpLEA5*, the late embryogenesis abundant protein gene from *Chimonanthus praecox*, possesses low temperature and osmotic resistances in prokaryote and eukaryotes. Int. J. Mol. Sci. 16, 26978–26990.

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.

Monfared, M.M., Simon, M.K., Meister, R.J., Roig-Villanova, I., Kooiker, M., Colombo, L., Fletcher, J.C., Gasser, C.S., 2011. Overlapping and antagonistic activities of *BASIC*

PENTACYSTEINE genes affect a range of developmental processes in Arabidopsis. Plant J. 66, 1020–1031.

Morran, S., Eini, O., Pyvovarenko, T., Parent, B., Singh, R., Ismagul, A., Eliby, S., Shirley, N., Langridge, P., Lopato, S., 2011. Improvement of stress tolerance of wheat and barley by modulation of expression of DREB/CBF factors. Plant Biotechnol. J. 9, 230–249.

Mukhopadhyay, M., Mondal, T.K., Chand, P.K., 2016. Biotechnological advances in tea (*Camellia sinensis* [L.] O. Kuntze): a review. Plant Cell Rep. 35, 255–287.

Muoki, R.C., Paul, A., Kumar, S., 2012. A shared response of *thaumatin like protein, chitinase*, and *late embryogenesis abundant protein3* to environmental stresses in tea [*Camellia sinensis* (L.) O. Kuntze]. Funct. Integr. Genom. 12, 565–571.

Paul, A., Singh, S., Sharma, S., Kumar, S., 2014. A stress-responsive *late embryogenesis abundant protein 7* (*CsLEA7*) of tea [*Camellia sinensis* (L.) O. Kuntze] encodes for a chaperone that imparts tolerance to *Escherichia coli* against stresses. Mol. Biol. Rep. 41, 7191–7200.

Pedrosa, A.M., Santos Martins, C.D.P., Goncalves, L.P., Cardoso Costa, M.G., 2015. Late embryogenesis abundant (LEA) constitutes a large and diverse family of proteins involved in development and abiotic stress responses in sweet orange (*Citrus sinensis* L. Osb.). PLoS One 10, e0145785.

Salleh, F.M., Evans, K., Goodall, B., Machin, H., Mowla, S.B., Mur, L.A., Runions, J., Theodoulou, F.L., Foyer, C.H., Rogers, H.J., 2012. A novel function for a redox-related LEA protein (*SAG21*/AtLEA5) in root development and biotic stress responses. Plant Cell Environ. 35, 418–429.

Shao, H.B., Liang, Z.S., Shao, M.A., 2005. LEA proteins in higher plants: structure, function, gene expression and regulation. Colloids Surfaces B Biointerfaces 45, 131–135.

Shih, M., Hoekstra, F.A., Hsing, Y.C., 2008. Late embryogenesis abundant proteins. In: Advances in Botanical Research, pp. 211–255.

Stacy, R.A., Aalen, R.B., 1998. Identification of sequence homology between the internal hydrophilic repeated motifs of group 1 late-embryogenesis-abundant proteins in plants and hydrophilic repeats of the general stress protein GsiB of *Bacillus subtilis*. Planta 206, 476–478.

Su, M., Huang, G., Zhang, Q., Wang, X., Li, C., Tao, Y., Zhang, S., Lai, J., Yang, C., Wang, Y., 2016. The LEA protein, ABR, is regulated by ABI5 and involved in dark-induced leaf senescence in *Arabidopsis thaliana*. Plant Sci. 247, 93–103.

Wang, M., Li, P., Li, C., Pan, Y., Jiang, X., Zhu, D., Zhao, Q., Yu, J., 2014. SiLEA14, a novel atypical LEA protein, confers abiotic stress resistance in foxtail millet. BMC Plant Biol. 14, 290.

Wang, W., Xin, H., Wang, M., Ma, Q., Wang, L., Kaleri, N.A., Wang, Y., Li, X., 2016. Transcriptomic analysis reveals the molecular mechanisms of drought-stress-induced decreases in *Camellia sinensis* leaf quality. Front. Plant Sci. 7, 385.

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

Wei, C., Yang, H., Wang, S., Zhao, J., Liu, C., Gao, L., Xia, E., Lu, Y., Tai, Y., She, G., Sun, J., Cao, H., Tong, W., Gao, Q., Li, Y., Deng, W., Jiang, X., Wang, W., Chen, Q., Zhang, S., Li, H., Wu, J., Wang, P., Li, P., Shi, C., Zheng, F., Jian, J., Huang, B., Shan, D., Shi, M., Fang, C., Yue, Y., Li, F., Li, D., Wei, S., Han, B., Jiang, C., Yin, Y., Xia, T., Zhang, Z., Bennetzen, J.L., Zhao, S., Wan, X., 2018. Draft genome sequence of *Camellia sinensis* var. *sinensis* provides insights into the evolution of the tea genome and tea quality. Proc. Natl. Acad. Sci. Unit. States Am. 115, E4151–E4158.

Wu, C., Hu, W., Yan, Y., Tie, W., Ding, Z., Guo, J., He, G., 2018. The late embryogenesis qbundant protein family in cassava (*Manihot esculenta* Crantz): genome-wide characterization and expression during abiotic stress. Molecules 23, 1196.

Xia, E.H., Zhang, H.B., Sheng, J., Li, K., Zhang, Q.J., Kim, C., Zhang, Y., Liu, Y., Zhu, T., Li, W., et al., 2017. The tea tree genome provides insights into tea flavor and independent evolution of caffeine biosynthesis. Mol. Plant 10, 866–877.

Yu, J., Lai, Y., Wu, X., Wu, G., Guo, C., 2016. Overexpression of *OsEm1* encoding a group I LEA protein confers enhanced drought tolerance in rice. Biochem. Biophys. Res. Commun. 478, 703–709.

Zeng, X., Ling, H., Yang, J., Li, Y., Guo, S., 2018. LEA proteins from *Gastrodia elata* enhance tolerance to low temperature stress in *Escherichia coli*. Gene 646, 136–142.

Zhao, P., Liu, F., Zheng, G., Liu, H., 2011. Group 3 late embryogenesis abundant protein in *Arabidopsis*: structure, regulation, and function. Acta Physiol. Plant. 33, 1063–1073.