Comparative Biochemistry and Physiology, Part B 160 (2011) 32–39

Contents lists available at ScienceDirect

Comparative Biochemistry and Physiology, Part B

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

Diverse *LEA* (late embryogenesis abundant) and *LEA-like* genes and their responses to hypersaline stress in post-diapause embryonic development of *Artemia franciscana*

Gang Wu [a,b], Hangxiao Zhang [a,b], Jing Sun [a,b], Fei Liu [a,b], Xiaomeng Ge [a], Wei-Hua Chen [a,c], Jun Yu [a,*], Weiwei Wang [a,d,**]

[a] CAS Key Laboratory of Genome Sciences and Information, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 100029, China
[b] Graduate University of Chinese Academy of Sciences, Beijing 100049, China
[c] European Molecular Biology Laboratory (EMBL), Heidelberg 69117, Germany
[d] Division of Gastroenterology, Department of Medicine, University of Alberta, Edmonton, Alberta, Canada T6G 2V2

ARTICLE INFO

Article history:
Received 29 March 2011
Received in revised form 16 May 2011
Accepted 16 May 2011
Available online 18 May 2011

Keywords:
Artemia franciscana
Late embryogenesis abundant
Hydrophilic protein
Post-diapause embryos
Hypersaline stress

ABSTRACT

From post-diapause cysts of *Artemia franciscana*, we defined fourteen *LEA* (late embryogenesis abundant) and *LEA-like* genes, including four novel members (*Afrlea1-5*, *Afrlea3-5*, *Afrlea3-like1* and *Afrlea3-like2*), which were classified into four groups: G1, G3, G3-like (LEA group3-like), and SMP-like (seed-maturation-protein-like), based on their conserved and diversified sequence motifs and amino acid compositions among bacteria, plants, and animals. We also validated six representative genes based on quantitative RT-PCR, including three *LEA* and two *LEA3-like* genes that are down-regulated when dehydrated cysts hatch to desiccation-intolerant larvae as well as one *SMP-like* gene that is slightly up-regulated. We further tested their responses to hypersaline stress for four representatives—one from each group—and found that the expression of *Afrlea1-5* and *Afrlea3-2* were inducible but not *Afrlea3-like1* and *Afrsmp-like*. This result suggested that the *LEA* and *LEA-like* genes may play different roles in resistance to hypersaline stress.

1. Introduction

Diapause, a seasonal pause in development, is programmed during an organism's life cycle and controlled by complex interplays of endogenous factors and various environmental cues. It is not a simple growth break but a complicated process that is divided into three stages: pre-diapause, diapause, and post-diapause quiescence (Kostal, 2006). *Artemia franciscana* is a primitive crustacean shrimp-like organism that lives in saline water and produces diapause embryos (known as cysts) whose development and metabolism are suspended (Anderson et al., 1970) and these cysts are able to remain viable for years without water (Clegg, 1997). Therefore, they often serve as an experimental model for studying animal desiccation process.

Water loss often leads serious consequences such as protein aggregation (Goyal et al., 2005), oxidative stress, higher cytoplasmic viscosity, and DNA damage (Caramelo and Iusem, 2009). One of the most important stress-associated genes is late embryogenesis abundant (LEA), discovered in many desiccation-tolerant organisms including plants, bacteria, and invertebrates (Tunnacliffe and Wise, 2007; Battaglia et al., 2008), and many of the *LEA* genes have been demonstrated to be relevant to water stress response (Welin et al., 1994; Maul et al., 1995; Gal et al., 2004). Interestingly, the nematode *LEA* gene expressed in a human cell line was reported to increase cell survival during dehydration (Chakrabortee et al., 2007).

Although significant progress has been made in studying *LEA* genes, we have not revealed their detailed molecular mechanisms about how they functionally interact with other cellular components. LEA proteins are hydrophilic, non-globular, and low-complexity, and capable of changing their conformations according to water content of the environment. For example, one group 3 or G3 LEA protein from nematodes was unfolded in its native state in solution but folded properly upon desiccation *in vitro* (Goyal et al., 2003). Another difficulty in understanding the function of *LEA* genes comes from their sequence divergence that results in the formation of large families, even within a single species. There are at least six different groups and hundreds of *LEA* genes widely distributed among organisms ranging from unicellular to multicellular organisms (Tunnacliffe and Wise, 2007; Battaglia et al., 2008). For instance, 5 *LEA-like* genes in *Escherichia coli* and 12 in *Saccharomyces cerevisiae* were predicted and validated under water deprivation (Garay-Arroyo et al., 2000). There are about 50 *LEA* genes found in *Arabidopsis thaliana* (Bies-Etheve et al., 2008; Hundertmark and Hincha, 2008). Therefore, it is necessary

* Correspondence to: J. Yu, CAS Key Laboratory of Genome Sciences and Information, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 100029, China. Tel.: +86 10 82995357; fax: +86 10 82995373.
** Correspondence to: W. Wang, Division of Gastroenterology, Department of Medicine, University of Alberta, Edmonton, Alberta, Canada T6G 2V2.
E-mail addresses: wugang@big.ac.cn (G. Wu), zhanghangxiao@big.ac.cn (H. Zhang), sunjing@big.ac.cn (J. Sun), liuf@big.ac.cn (F. Liu), gexiaomeng@big.ac.cn (X. Ge), weihua.chen@embl.de (W.-H. Chen), junyu@big.ac.cn (J. Yu), wangweiwei@big.ac.cn (W. Wang).

1096-4959/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.cbpb.2011.05.005

to study *LEA* genes in each species, which are appropriately stratified by physiochemical characteristics, sequence homology, and functional classification. Moreover, different groups of LEA proteins are hypothesized to play distinct cellular roles. For example, G1 LEA proteins in *A. franciscana* are mostly cytosol-localized with some in mitochondria, even can translocate into nuclei under anoxic conditions (Warner et al., 2010). G3 LEA proteins are widely localized in cytoplasm, mitochondria, chloroplast, and even located in the endoplasmic reticulum in plants (Tunnacliffe and Wise, 2007).

In this study, based on proteomic and transcriptomic discoveries of *A. franciscana* genes in developing post-diapause embryos (Wang et al., 2007; Chen et al., 2009), we defined four new members and reclassified ten known *LEA* and *LEA-like* genes previously reported in *A. franciscana* into four groups (Hand et al., 2007; Wang et al., 2007; Chen et al., 2009; Menze et al., 2009; Sharon et al., 2009; Warner et al., 2010). We also studied the function of different *LEA* and *LEA-like* genes in responding to hypersaline stress in developing *A. franciscana* embryos.

2. Materials and methods

2.1. Identification and analysis of A. franciscana LEA and LEA-like genes

We downloaded all LEA-related protein sequences from NCBI to form a dataset, including ten reported members in *A. franciscana* and compared them to 8,018 unigene clusters generated in our lab previously (Chen et al., 2009). The parameters for blastx were, *E*-value<0.001, identity>25%, alignment length>40 amino acids. Open reading frames were subsequently identified and analyzed based on Clustalx (Thompson et al., 1997). We used ProtParam to organize sequence information and Kyte–Doolittle algorithm (Kyte and Doolittle, 1982) for hydropathy analysis (ProtScale and a 9-animo-acid-residue sliding window). We also used RADAR (http://www.ebi.ac.uk/Tools/Radar/index.html) and TRUST(http://www.ibi.vu.nl/programs/trustwww) for detecting sequence repeat motif, GOR IV and SOPMA for secondary structure predictions (http://au.expasy.org/tools), and TargetP (http://www.cbs.dtu.dk/services/TargetP) and MitoProt (http://ihg.gsf.de/ihg/mitoprot.html) for intracellular localization signal analysis. We used blastn (*E*-value<1e-10, identity>90%, alignment length>60 bp) to search for EST copy numbers for LEA candidates in cDNA libraries (Chen et al., 2009). Phylogenetic trees were built by using Mega (Kumar et al., 2008) and based on the neighbor-joining method with bootstrap tests of 1000 replicates. The repeat motif of each LEA and LEA-like protein sequences were displayed with MEME, WeboLogo or ICM-Pro.

2.2. Classification of A. franciscana LEA and LEA-like genes

We classified our newly-discovered and previously-known (from the literature and public databases) *A. franciscana LEA* and *LEA-like* genes according to the Wise rule (Wise, 2003) and renamed them for convenience. For example, *Afrlea3-1* means the first *A. franciscana LEA* gene defined in G3, which was originally named as *Afrlea1* (Hand et al., 2007). The *LEA* and *LEA-like* genes were finally distinguished based on their expression patterns from dehydrated cysts to larvae.

2.3. Hypersaline stress treatment

The dehydrated cysts were collected from the Bohai Bay, Northern China, and identified as *A. franciscana* based on DNA barcoding (Fig. A.1; Wang et al., 2008). We evaluated hatching rate (Fig. A.2) in a 48 h time window in 2.5% sea salt at 27 °C with modest illumination, following an improved procedure described previously (Clegg, 2005). The samples were collected at 3 h, 10 h, and 48 h after incubation. The dehydrated cyst was treated as the starting point (0 h). In hypersaline stress experiments, we used NaCl solutions instead of seawater in order to exclude the influence of unknown ions and soluble molecules. We treated the cysts with 5% or 10% NaCl for 0.5 h and 1.5 h at 9.5 h after incubation at 27 °C.

2.4. RNA extraction and RT-qPCR

We extracted total RNA from *A. franciscana* cysts with Trizol (Invitrogen) according to the manufacturer's instruction and removed residue DNA with RNase-free DNaseI (New England Biolabs). For RT-qPCR, we used 4 μg total RNA for reverse transcription and primed with poly(T) oligonucleotides (M-MuLV reverse transcriptase, New England Biolabs). RT-qPCR was performed in an ABI PRISM 7500 system (Life Technology, Foster City, CA) and sequence-specific primers were designed by using Oligo 6.0 (Table A.1; http://www.oligo.net). Data were acquired under the same reaction condition (95 °C for 2 min; 40 cycles of the 3-staged temperature profiles of 95 °C for 15 s, 56 °C for 20 s, 68 °C for 40 s). Melting curves for each PCR experiment were carefully monitored to avoid nonspecific amplifications. Experiments from sample preparation to RT-qPCR were replicated three times. In each replicate experiment, 100 mg cysts were used for each sample, and in RT-qPCR normalization was firstly performed by manufacturer's software with reference dye in at least three nested replicates. Gene expression levels were transformed by using comparative Ct formula (Schmittgen and Livak, 2008) by further normalization based on the expression of α-tubulin. We calculated mean and standard error of the mean values for fold change with three replicate results. Student's *t*-test was used for comparing dehydrated cysts and larvae development stages. Paired *t*-test was used for analyzing the NaCl treated and control samples with *p*-value correction with Benjamini method (Benjamini and Hochberg, 1995).

3. Results

3.1. The characterization of A. franciscana LEA and LEA-like genes

We defined fourteen *LEA* and *LEA-like* genes based on intensive literature search and our own data containing 8,018 unigenes (Text A.1). Among them, there were four novel members: one in G1 or *Afrlea1-5*, one in G3 or *Afrlea3-5*, and two in LEA3-like, *Afrlea3-like1* and *Afrlea3-like2* (Table 1). According to the Kyte–Doolittle hydropathy plot, we found that the LEA proteins (Fig. 1A to D) were more hydrophilic than the LEA-like proteins (Fig. 1E and F), and G1 proteins (Fig. 1A and B) were more hydrophilic than those of G3 (Fig. 1C and D).

Since LEA proteins have rather characteristic amino acid compositions, we also examined the amino acid usage of *A. franciscana* LEA and LEA-like proteins; they are relatively low in C, W, F and Y, but very high in A, R, E, Q and G, similar to that of classical LEA and SMP proteins (Fig. A.3). G1 LEA proteins appeared to have distinct amino acid usage but not that of G3 LEA proteins which are similar to LEA3-like and SMP-like proteins. All fourteen *A. franciscana* protein sequences shared a repeat motif in 1 to 7 units (Fig. 2). Our phylogenetic analysis revealed that LEA proteins in animals tended to cluster together at intraspecies level, but display divergence at interspecies level (Fig. A.4).

3.2. Distinct characteristics of different A. franciscana LEA groups

The proteins encoded by six G1 genes shared a 20-amino acid, highly conserved motif -RAEQLGHEGYQEMGQKGGQA/T, which was usually found in plants. *A. franciscana* is the first animal reported to possess this motif in its LEA (Sharon et al., 2009). We also searched for possible motif among all G1 LEA proteins, and selected 27 motifs which were repeated at least 3 times in a single sequence (Text A.2). Analyzing these motifs, we found that the complete versions of these motifs are dividable into two sub-motifs. For instance, the sub-motif E-Y-E-G-KGG is conserved among archaea, bacteria, and eukaryotes, whereas other sub-motifs are