LEA proteins in higher plants: Structure, function, gene expression and regulation

Shao Hong-Bo a,b,c,*, Liang Zong-Suo a, b, d, Shao Ming-An a, e

a State Key Laboratory of Soil Erosion and Dryland Farming, The Centre of Soil and Water Conservation & Ecosystem Research, Chinese Academy of Sciences and Northwestern Sci-tech University of Agriculture and Forestry, Yangling 712100, China
b Molecular Biology Laboratory, Bioinformatics College, Chongqing University of Posts & Telecommunications, Chongqing 400065, China
c Biological Laboratory, Chemical Engineering College, Qingdao University of Science and Technology, Qingdao 266042, China
d Life Science College, Northwestern Sci-tech University of Agriculture and Forestry, Yangling 712100, China
e Institute of Graphical Sciences and Resource, Chinese Academy of Sciences, Beijing 100101, China

Received 9 May 2005; accepted 28 July 2005

Abstract

Late embryogenesis abundant (LEA) proteins are mainly low molecular weight (10–30 kDa) proteins, which are involved in protecting higher plants from damage caused by environmental stresses, especially drought (dehydration). These findings and the fact that the breeding of drought tolerant varieties would be of great value in agriculture, form the basis of search for anti-drought inducible genes and their characterization. LEA proteins are generally classified into six groups (families) according to their amino acid sequence and corresponding mRNA homology, which are basically localized in cytoplasm and nuclear region. LEA protein synthesis, expression and biological activities are regulated by many factors (e.g. developmental stages, hormones, ion change and dehydration), signal transduction pathways and lea genes. No tissue-specific lea gene expression has been considered as one main regulatory mechanism on the basis of extensive studies with the model plant, Arabidopsis thaliana. The study of the regulatory mechanism of lea gene expression is an important feature of modern plant molecular biology.

© 2005 Elsevier B.V. All rights reserved.

Keywords: Seed plants; LEA proteins; Gene structure; Function and expression; Drought resistance

1. Introduction

Higher plants dominate terrestrial ecosystems and play most important economic and social roles in human life [1,2], which contribute to their strong adaptive ability owing to a long period of evolution under the pressure of natural and human selection [3,4]. During such time span, higher plants have developed multi-pathway, multi-level and multi-scale survival strategies for continual changes in the environment, which include anatomical, physiological, biophysical, biochemical, genetic, developmental and reproductive biological changes in response to adverse conditions [5]. The evolution of LEA proteins is one of these changes, which plays an important role in resistance to drought. This implies that studies on tolerant proteins, and the isolation, identification and functional analysis of their genes will be of great benefit to the breeding of drought-resistant crops [6,7,47–52]. It is common knowledge that drought related food shortages is a worldwide problem. Especially in sub-Saharan Africa [8,9,47–52]. Drought is one of the main factors limiting crop production [10–16]. Research on the biology and genetics of drought resistance and breeding of drought-resistant crops have, thus, become important field of contemporary research in plant molecular biology and molecular breeding. These efforts have been greatly aided by the sequencing of the Arabidopsis thaliana genome [6,17,18]. Here, we discuss the role of LEA protein structure, function and gene expression in the development of drought-resistant crops.
2. LEA protein distribution and types in higher plants

LEA proteins are formed during the late period of seed development accompanied by dehydration. They are proteins with small molecular weights, ranging mainly from 10 to 20 kDa and above 30 kDa [19]. Dure and Croul [17] first studied LEA proteins in developing cotton seeds. Subsequently, researchers detected their existence in wheat, barley, maize, rice, sunflower, potato, grape, apple, bean, Arabidopsis, tomato, rye, soybean, carrot and so on [19–28]. LEA proteins exist mainly in higher plant seeds, but have also been found in seedlings, roots and other organs [29]. LEA proteins are mainly localized in cytoplasm and nuclear regions [26,29].

With the advent of new detection techniques [30], it has become possible to classify LEA proteins based on molecular hybridization (Northern hybridization, Western blotting) and immunological methods. Based on their amino acid sequence and mRNA homology, LEA proteins are basically divided into five groups, Group 1–5 (Dure, 1993; Ingram and Bartels [40]; He and Fu [19]; Zhang and Zhao [20]).

3. LEA protein structure

LEA proteins in higher plants are mainly composed of hydrophilic amino acids ordered in repeated sequence (e.g. Gly and Lys), forming hyper-hydrophilic and thermal stability. Advanced structure of such protein contains non-periodic linear and α-helical structure without thermal dominant state and corresponding dehydrated proteins exist in a natural form of dimers.

Group 1 proteins (such as D19) have a conservative sequence made up of 20 amino acid residues in repeated copies, which can absorb a large amount of bound water [24]. Group 1 protein genes play a role in endosperm development and osmotic protection of vegetative organs in higher plants (Clark and Marcotte, 1999). Group 2 proteins (e.g. D11) include a 15-amino acid conservative structure (EKKMR-DKKEKLP) at the C-terminus, and play an important role in metabolism as molecular chaperones and defending protein structure with a close connection to higher plant resistance to drought [25,31,32]. Group 3 proteins (e.g. D7) are made up of 11 amino acids (TAQAAKEKAGE) in 13 repeats, resulting in amphipathic α-helical structure, and are involved in enriching ions during dehydration of higher plants (Zhang and Zhao [20]; Yu [21]). Barley PMA1949, carrot Dc8, and soybean pGmPM2 proteins belong to this family [22,29,33]. Group 4 proteins include cotton LEA4, D113 and Cucurbita maxima PGC27-45, which are devoid of repeated motif sequences and contain a conservative region at N-end to form amphipathic α-helical structure. The above proteins can further form a subsidiary structure adaptive to conformational changes of other proteins and function in protecting membrane stability and integration during drying and dehydration [34,35].

4. LEA protein functions

LEA proteins mainly play functions in dehydration tolerance and storage of seeds and in whole-plant stress resistance to drought, salt, and cold. Farrant et al. (1992) experimented with the recalcitrant seeds of Avicennia marina and Podocarpus henkelii and concluded that one of the important reasons was the absence of LEA proteins in dehydration sensitive seeds. Further study showed that LEA proteins are expressed through all the developmental stages with different expression levels and no tissue specificity. For instance, Em, RAB21 and dehydrins in seeds can be found in the root, stem, leaf, callus and suspension cultures of higher plants under ABA or NaCl induction [21,28,36]. In cotton seedlings, in vitro treatment could lead to the accumulation of LEA protein mRNAs [22,29,31,34,37]. Almonguera et al. (1992) reported that LEA mRNAs and heat shock mRNAs accumulated in a concerted way in sunflowers. Zhang et al. (2000) applied yeast expression systems to investigate the effect of LEA proteins on cellular metabolism under different water stresses and found that transfer of two genes, tomato le4 (Group 2) and barley HVA1 (Group 3) into such systems could promote earlier expression of HVA1 gene in the medium containing 1.2 mol L−1 NaCl under GAL1 promoter, increase the expression of both genes in 1.2 mol L−1 KCl medium, resulting in growth reduction. These results imply that LEA proteins play a special role in protecting cytoplasm from dehydration. Functions related to other aspects remains unknown.

5. Gene expression and regulation of LEA proteins in higher plants

LEA protein gene expression in terms of time course starts from the late period of maturation and initiation period of drying reaches its peak in progressive dehydration and sharply decreases after some hours of germination [19,23,27,38].

Many reports show that LEA protein gene expression has no tissue-specificity at the levels of tissues and organs as the gene can express in cotyledons, panicles of seeds [26] and also in stems, leaves and roots (vegetative tissues) [16,18,28,36].

Because of different classification methods for developmental stages in diverse plants, there exist some controversial results (Galau et al., 1991). Results from our lab are in support of Hughes and Galau (1989) who used barley seeds and in vitro seedlings as experimental materials at four defined periods of development (cotyledon, seedling, maturation and mature termination) [38,39,51].
Higher plant adaptation to changing environment has many mechanisms involved at different levels, one of which is LEA protein gene expression and regulation formed during long evolutionary history of natural selection and artificial selection. The characteristics, which are perceived in anatomy, physiology, biochemistry, biophysics and development biology, are mainly the responses to related gene expression and de-regulation at the molecular level. Up to now, it is difficult to describe a detailed clear picture concerning this aspect [25,33,38–52]. In fact, various factors and conditions and processes influence LEA protein gene expression, among which ABA is considered the most important, especially in reducing the harm caused by drought and is connected directly or indirectly with other regulatory circuits [Shao et al. [12,46]; Somerville and Dangl [1]; Chaves et al. [35]]. Information from the model plant A. thaliana has shown that there is a universal signal transduction network system in higher plants, on the basis of and through which other factors produce different effects, such as ABA concentration changes and drought as signal stimuli [12,34,38,39,47–50]. On the other hand, there is a stress-threshold problem (stress intensity), over-maximum stress degree will lead to initiation and cascade reactions of other gene expression pathways like related senescence and death gene expression reactions (Zhu, 2001; Xiong et al. [45]; Bray [47]; Shao et al. [12,46]). So, four steps at least are involved in LEA protein gene expression and regulation induced by drought: signal recognition, signal transduction, signal amplification and integration, LEA protein gene expression responses and its product formation. We will take ABA, which has recently been shown to be involved clearly in inducing LEA protein gene expression, as an example for further review. Concerning this aspect, there are at least three different induced pathways: ABA-dependent type; ABA-induced type; ABA-irresponsive type [27,40–42,52].

ABA-dependent type gene expression is determined by endogenous ABA accumulation or exogenous ABA levels. Under drought stress, the endogenous ABA level increases. In barley seedlings and beans, drought can result in high levels of ABA by 57–160 times compared with those of controls [36,40]. ABA physiological functions are realized by corresponding gene expression. Many genes induced by ABA have been isolated, cloned, identified, and subjected to functional analysis (Bray [47]; Chaves et al. [35]). It has been proved that cis-acting and trans-acting factors are related to ABA-responsive gene transcriptional regulation. There are some differences in the DNA sequence containing ABRE, in which there exists a G-box made of 6CT sequence, and it can bind the transcriptional factor, bZIP [Shen et al. (1995) reported that CEI-binding elements needed ABA-responsiveness under the promoter of barley HVA22. In A. thaliana, there are at least 58 genes coding bZIP [36,40,41]. Many of LEA protein genes can be induced by ABA. A. thaliana drought-induced gene RD22 is regulated by ABA [42]. Under severe stress, two DNA elements (MYC-like element and MYB-like element) were identified, which were related to ABA-induced gene expression. The DNA-binding protein gene rd22 B1P1, coding MYC-like elements, can be induced by drought and ABA [43]. Flores et al. [44] studied Phaseolus vulgaris by treatments of ABA and water stress, induced 6 cDNA clones, in which there were two types of LEA protein genes [44]. Under the water pressure of 0.35 MPa (16 h), these genes reached their expression peak with a decrease in the progressive stress and after recovery of water: these genes could obtain transient expression with the above level. These results showed that short-term and rapid treatment cannot lead to changes in the expression of these genes.

In higher plant adaptation to drought, mRNA accumulation and activities can respond to gene expression to large extent. ABRE is related to water stresses, which contains CACGTG and G-box. Wheat Emilia and rice fabA6A gene promoter elements are ABRES, combined with 35S promoter gene in many copies and regulated by ABA [42,44]. Iuchi et al. [43] induced 10 cDNA clones from a high-tolerance-drought cowpea. Northern hybridization implied that higher salt stress could result in the expression of three genes, cPRD8, cPRD14, cPRD22, but not cold or heat stress. In the defined range of ABA applied exogenously, cPRD and cPRD22 could be induced, but not cPRD14, which suggested that there were two signal transduction pathways between water stress and the expression of cPRD genes in cowpea.

By studying A. thaliana ABA-insensitive mutants (aba mutants), some of the genes can be induced by drought, salt, and cold, so it is considered that the expression of these genes do not need ABA under cold or drought [45]. Some of these genes, such as rd29A, lhn78 and cxr78 have been analyzed under drought-inducing conditions [46–52]. During drying and cold stress, there are at least two independent systems, by which gene expression can be regulated, i.e. ABA-dependent and non-ABA-dependent type. rd29A is a non-ABA-dependent cis-acting element [Yamaguchi-Shinozaki et al., 1993, 1994]. By fusing rd29A promoter with GUS reporter gene and transferring the combined construct into A. thaliana and tobacco plants, dehydration, low temperature or higher salt conditions could produce obvious expression of the constructed gene under such promoter. Further analysis demonstrated that the promoter region of rd29A contained a conservative sequence of 9 bp (TACCGACAT, DRE), which is necessary for rd29A expression. It is easy to know that dehydration, low temperature or higher salt can each induce the genes with DRE cis-acting elements, but cannot induce the genes with ABRE.

6. Summary and perspective

Over the past 10 years, many advances have taken place in this challenging field. Although many molecular aspects are
from the model plant A. thaliana, these have indeed deepened our understanding of higher plant resistance to drought. This anti-drought character is a quantitative character controlled by many genes and impacted by a large number of environmental, anatomical, physiological, biophysical, biochemical and developmental factors. The characteristic result we observe under a given condition is practically a comprehensive character. Many related genes (cDNA clones) have been isolated and sequenced, but the functions and biochemical properties of the proteins coded by these genes depend upon predicting according to public database resources, and remain unclear. Their space structure is also a challenge. Besides the evolutionary relationship of different LEA protein groups in diverse higher plant species with their adaptation to adverse environment needs further effort, so the range of the tested materials should be enlarged to explore the universal principle in functions, structure, gene expression, and evolution. All these beg answers.

Information from LEA protein gene expression shows that it is a multi-functional stress protein to maintain normal metabolism for higher plants under severe conditions, and its gene can code RNA-regulatory-proteins, which can regulate related events, such as gene expression and development. The mechanism is unknown (Wu et al., 2003). In addition, what is the stress threshold (e.g. drought degree)? How much drought intensity can prime LEA protein gene expression? Which is the signal transduction pathway? What is the cross-talk between this pathway and others? How many components are involved in transcriptional and translational level regulation, respectively? How to integrate the related data into a clear picture and use it in practice for oriented-biotechnological anti-drought breeding is another big challenge.

Up to date, there have been mainly two ways to study LEA protein genes: select mutant genes from model plants and analyze stress responses of crops. The extensive application of transgenic plants and microarrays is one developing direction. The related evidence directly from transgenic plants and microarrays is one development. The characteristic result we observe under a given condition is practically a comprehensive character. Many related genes (cDNA clones) have been isolated and sequenced, but the functions and biochemical properties of the proteins coded by these genes depend upon predicting according to public database resources, and remain unclear. Their space structure is also a challenge. Besides the evolutionary relationship of different LEA protein groups in diverse higher plant species with their adaptation to adverse environment needs further effort, so the range of the tested materials should be enlarged to explore the universal principle in functions, structure, gene expression, and evolution. All these beg answers.

Information from LEA protein gene expression shows that it is a multi-functional stress protein to maintain normal metabolism for higher plants under severe conditions, and its gene can code RNA-regulatory-proteins, which can regulate related events, such as gene expression and development. The mechanism is unknown (Wu et al., 2003). In addition, what is the stress threshold (e.g. drought degree)? How much drought intensity can prime LEA protein gene expression? Which is the signal transduction pathway? What is the cross-talk between this pathway and others? How many components are involved in transcriptional and translational level regulation, respectively? How to integrate the related data into a clear picture and use it in practice for oriented-biotechnological anti-drought breeding is another big challenge.

Acknowledgements

This work is jointly supported by State Key Basic Research Development Plan of China (G2000018605), Major Research Program of NSFC (90102012), Knowledge Innovation Engineering of the Chinese Academy of Sciences (KZCX3-06-2-A), Shao MA-Innovation Group Program Project of Northwestern Sci-tech University of Agriculture and Forestry, Doctoral Initiation Foundation of Chongqing University of Posts & Telecommunications and West-China Key Research Project of National Natural Science Foundation of China (90302005). Many thanks are given Professor Adam Wilkins in Cambridge University, UK and Professor I.K. Vaid in University of Florida, USA, for their critical reading of the manuscript and providing many instructive suggestions.

References