การวิเคราะห์ลำดับนิวคลีโอไทด์ของยีนทนแล้ง (*EgLEA*4) จากปาล์มน้ำมัน (*Elaeis guineensis* Jacq)

Sequence Analysis of Drought-tolerance (*EgLEA*4) Gene from Oil Palm (*Elaeis* guineensis Jacq)

วลีรัตน์ ทองอินทร์*, อลิษา หนักแก้ว, วราพร วรรณนา และ อมรรัตน์ พงศ์ดารา

Waleerath Thongin*, Alisa Nakkaew, Warapond Wanna and Amornrat Phongdara

สถานวิจัยจีโนมและชีวสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยสงขลานครินทร์ หาดใหญ่ สงขลา 90110

Center for Genomics and Bioinformatics Research, Faculty of Science, Prince of Songkla University, Songkhla 90110

**Correspondence author: yaibong_fgx@hotmail.com*

บทคัดย่อ

Late embryogenesis abundant (LEA) เป็นโปรตีนที่เกี่ยวข้องกับสภาวะความเครียดโดยเฉพาะการขาดน้ำของพืช จากการวิเคราะห์ ยีน *EgLEA*4 ที่ได้มาจากการทำห้องสมุด EST ของปาล์มน้ำมัน (*Elaeis guineensis* Jacq) พบว่า ยีนโครงสร้างของ *EgLEA*4 มีขนาด 1505 คู่เบส ประกอบด้วย 2 เอกซอน (108 และ 378 คู่เบส), 1 อินทรอน (313 คู่เบส) และบริเวณโปรโมเตอร์ของ *EgLEA*4 (447 คู่เบส) ลำดับนิวคลีโอไทด์ที่ไม่มีการแปลรหัสด้านปลาย 5' แสดงลำดับนิวคลีโอไทด์ของโปรโมเตอร์ เช่น ABA-responsive element (ABRE) และ salt/drought/dehydration responsive element (DRE) โปรตีนที่ได้เป็นส่วนที่ชอบน้ำ 50.3 เปอร์เซ็นต์ และเป็นส่วนที่ไม่ชอบน้ำ 49.7 เปอร์เซ็นต์ และจากการวิเคราะห์ด้วย Hydropathicity ทำนายว่าโปรตีนนี้เป็นโปรตีนที่ชอบน้ำ รูปแบบโมทีฟของยีน *EgLEA*4 ประกอบด้วย 5 โมทีฟ (2, 1, 3, 4 และ 5) ยีน *EgLEA*4 มีการแสดงออกสูงในเนื้อเยื่อชั้นกลางของผลปาล์มและมีการแสดงออกสูงสุดในระยะท้ายของการพัฒนาผลปาล์ม เนื่องจากปริมาณน้ำในเนื้อเยื่อผลปาล์มมีค่าน้อยลง

ABSTRACT

The late embryogenesis abundant (LEA) proteins in oil palm have been implicated as part of the stress response especially those associated with drought in plants. We first isolated the *EgLEA*4 gene from an EST library of oil palm (*Elaeis guineensis* Jacq.). The structural gene of *EgLEA*4 consisted of 1505 bp containing 2 exons (108 and 378 bp), 1 intron (313 bp), and the promoter region of an *EgLEA*4 (447 bp). The 5' untranslated region showed a putative promoter sequence such as an ABA-responsive element (ABRE) and a salt/drought/dehydration responsive element (DRE). The protein sequence analysis showed that it consisted of 50.3% hydrophilic residues and 49.7% hydrophobic residues. Hydropathicity analysis predicted that this protein was a hydrophilic protein. The amino acid patterns of *EgLEA*4 consisted of five conserved motifs (II, I, III, IV and V). The high expression of the *EgLEA*4 gene in the mesocarp tissue and the highest expression in the late stage of oil palm fruits development were associated with the decreased water content of the fruit's tissue.

คำสำคัญ: โปรตีน late embryogenesis abundant, การขาดน้ำ, ปาล์มน้ำมัน, *EgLEA*4

Keywords: late embryogenesis abundant protein, drought, oil palm, *EgLEA*4

INTRODUCTION

In plant, the late embryogenesis abundant (LEA) proteins have been claimed to be involved in the plant's water stress response caused by drought and/or cold stress (Goyal et al., 2005). In a study of Arabidopsis, a group-4 LEA protein was shown to contribute the ability of the plant that respond to drought stress. This protein accumulated during the late stages of seed maturation, but was not found during the seed germination. It was assumed that, because of water deficits only occurred during the later stages of seed maturation (Bies-Etheve et al., 2008; Galau et al., 1987). Group 4 of the *LEA* proteins have hydrophilic features and contain a high percentage of small amino acid residues such as glycine, alanine, and serine (Battaglia et al., 2008). LEA4 proteins have conserved regions near to the N-terminus, that cover about 70-80 amino acid residues, compared to the less conserved regions near the C-terminus (Dure, 1993). Furthermore, an *in vitro* study of the LEA4 family of proteins of Arabidopsis has shown that they were present during controlled dehydration experiments, and also prevented the inactivation of LDH, even after a 99% water loss (Reyes et al., 2005). In this work, we have provided more information that may help to obtain a clearer understanding on the role of the LEA4 proteins. This includes information about their sequence conservation, variation, and as well as their expression patterns during the developmental stages of the friut.

MATERIALS AND METHODS

Total RNA was extracted from the mesocarp of oil palm fruit during their 22 weeks of development. Semi-quantitative RT-PCR was performed. The resultant chromosomal DNA was extracted and used as the template for genome walking and structural gene analysis of the *EgLEA*4. The PCR product was separated by electrophoresis on a 1.5% (w/v) agarose gel, and a semi-quantitative analysis was carried out using the Quantity One program and the SPSS program. The obtained PCR product was ligated to the pGEM®-T Easy vector (Promega) and transformed into *Escherichia coli* as described by Maniatis et al., 1982. The recombinant plasmids were then purified and sequenced using the Applied BioSystem 377 Sequencer (Perkim Elemer). Finally, the 5' untranslated region of the *EgLEA*4 was analyzed using the Bioinformatic tools and the Genomatix software.

RESULTS AND DISSCUSSION

Using semi-quantitative RT-PCR for analysis the expression of the *EgLEA*4 gene during the development stages of the mesocarp compared to the expression of the *18s rRNA* as the internal control, the *EgLEA*4 gene achieved its highest expression during the late stage of development of the oil palm fruits (22 weeks) (Figure1). This correlated with the water deficit stage. A similar result had been found for the expression of the *AhLEA*4 gene in *Arachis hypogaea* that was only detected in their seeds (Su et al., 2011). The amino acid sequence alignment of different plant LEA4 proteins showed that all had a similar five conserved patterns of the motifs II, I, III, IV, and V. This confirmed that the EgLEA4 was part of the group 4 LEA proteins (Figure 2). Therefore, the structural gene of *EgLEA*4 consisted of 1505 bp containing 2 exons (108 bp and 378 bp), 1 intron (313 bp), and the promoter region of an *EgLEA*4 protein (447 bp). This result is similar to the *GhLEA*4 (accession number: M73752) that has two exons (193 bp and 522 bp) and an intron of 101 bp (Galau et al.,

1993). Furthermore, the EgLEA4 promoter sequence at the 5'-untranslated region was obtained by the GenomeWalking™ method and analyzed by the Genomatix software (MatInspector database). There were two putative cis-acting regulatory elements perhaps involved in the abiotic stress response, such as an ABA-responsive element (ABRE) and a salt/drought/dehydration responsive element (DRE) detected (Figure 3).

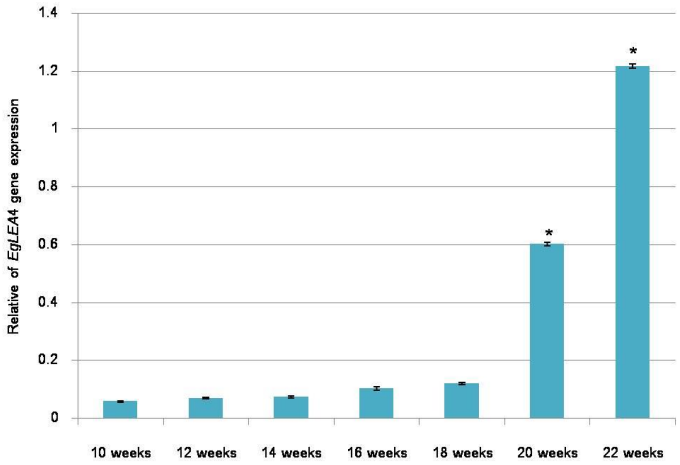

Figure 1 Relative *EgLEA4* gene expression from different stages of the oil palm fruits development. Bars show the relative expression of the *EgLEA4* gene compare to the expression of the *18s rRNA* gene as the internal control. The highest expression was observed at 20 and 22 weeks of fruit development, and they were significantly (P<0.05) higher when compared to all previous weeks. Values represent the means and s.e.m (N=3).

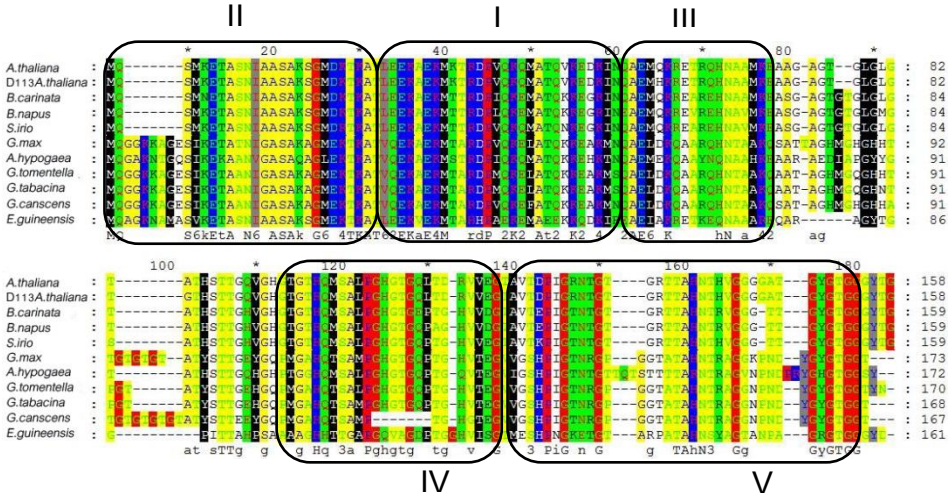

Figure 2 Multiple amino acid sequence alignment of the EgLEA4 protein from oil palm (*Elaeis guineensis* Jacq; AFG26321.1) and the LEA4 protein from other various plant species; *A.thaliana* (*Arabidopsis thaliana*; NP_196294.1), D113*A.thaliana* (LEAD113 *A.thaliana*; CAA63008.1), *B.carinata* (*Brassica carinata*; AAT77224.1), *B.napus* (*B.napus*; AAT77223.1), *S.irio* (*Sisybrium irio*; AAY26119.1), *G.max* (*Glycine max*; AAG37440.1), *A.hypogaea* (*Arachis hypogaea*; ADQ91840.1), *G.tomentella* (*G.tomentella*; AAG37451.1), *G.tabacina* (*G.tabacina*; AAG37441.1) and *G.canscens* (*G.canscens*; AAG37439.1). The enclosed black boxes indicate the five conserved motifs.

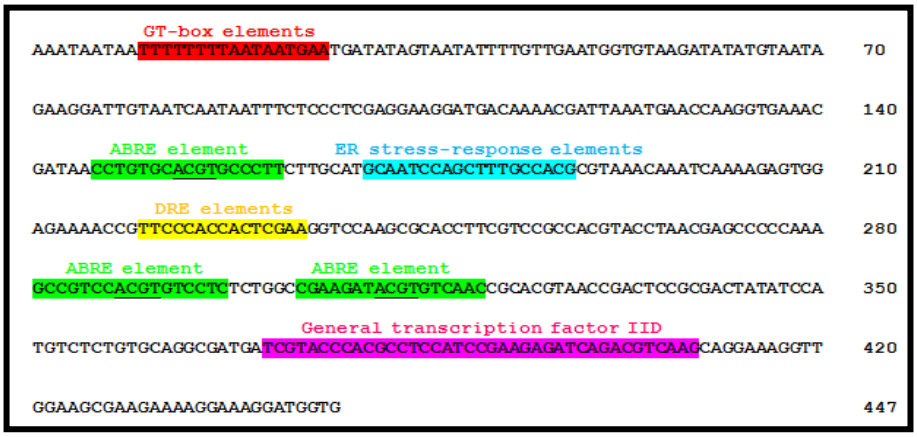

Figure 3 Positions of the various recognition elements from the promoter sequence of the nucleotides of the 5' untranslated *EgLEA4* gene in oil palm.

CONCLUSION

In this study, the molecular information, sequence analysis, gene structure and expression have provided some basic knowledge about the *EgLEA4* gene which may be useful for the improvement of drought tolerance in oil palm and other crops.

ACKNOWLEDGEMENTS

This research was supported by the Research Assistant (RA) funded by the faculty of Science (1-2554-02-007) to Ms.Waleerath Thongin and the National Research University Project of Thailand's Office of the Higher Education Commission for financial support, the Graduated School, the Center for Genomics and Bioinformatics Research, the Faculty of Science, and Prince of Songkla University. We thank Dr. Brian Hodgson for assistance with the English and valuable comments.

REFERENCES

Battaglia M, Olvera-Carrillo Y, Garciarrubio A, Campos F, Covarrubias AA. The enigmatic LEA proteins and other hydrophilins. Plant Physiology. 2008; 148(1): 6-24.

Bies-Etheve N, Gaubier-Comella P, Debures A, Lasserre E, Jobet E, Raynal M, Cooke R,

Delseny M. Inventory, evolution and expression profiling diversity of the LEA (late Embryogenesis abundant) protein gene family in *Arabidopsis thaliana*. Plant Molecular Biology. 2008; 67(1-2): 107-124.

Dure L. Structural motifs in LEA proteins In TJ Close, EA Bray, eds, Plant Responses to Cellular Dehydration during Environmental Stress. America Society of Plant Physiologists, Rockville. 1993; 91-103.

Galau GA, Bijaisodarat N, Hughes DW. Accumulation kinetics of cotton late embryogenesis-abundant mRNAs and storage protein mRNAs: coordinate regulation during embryogenesis and the role of abscisic acid. Development Biology. 1987; 123(1): 198-212.

Galau GA, Wang HYC, Hughes W. Cotton *Lea5* and *LEA14* encode atypical late embryogenesis-abundant protein. Plant Gene Register. 1993; 101: 695-6.

Goyal K, Walton LJ, Tunnacliffe A. LEA proteins prevent protein aggregation due to water stress. Biochemistry Journal. 2005; 388: 151-7.

Maniatis T, Fritsch EF, Sambrook J. Molecular cloning a laboratory manual. Cold Spring Harbor Laboratory. 1982; 11(2).

Reyes JL, Rodrigo MJ, Colmenero-Flores JM, Gil JV, Garay-Arroyo A, Campos F, Salamini F, Bartels D, Covarrubias AA. Hydrophilins from distant organisms can protect enzymatic activities from water limitation effect *in vitro*. Plant, Cell & Environment. 2005; 28: 709-718.

Su L, Zhao CZ, Bi YP, Wan SB, Xia H, Wang XJ. Isolation and expression analysis of LEA genes in peanut (*Arachis hypogaea* L). Journal of Bioscience. 2011; 36(2): 223-8.