Kastamonu Uni., Orman Fakültesi Dergisi, 2019, 19 (3): 299-309 Research Article
Kastamonu Univ., Journal of Forestry Faculty
Doi:10.17475/kastorman.662581

Identification and Characterization of LEA Genes in Ash Tree *(Fraxinus excelsior)* Genome

Aslı UĞURLU BAYARSLAN

Kastamonu University, Faculty of Arts and Sciences, Department of Biology, Kastamonu, TURKEY
augurlu@kastamonu.edu.tr

Received Date: 29.01.2019 Accepted Date: 23.08.2019

Abstract

Aim of study: LEA proteins have a critical role in the abiotic stress response of plants. Ash belongs to the olive family and its genome sequence is complete. The *LEA* genes has not been extensively analyzed, although there are annotations in the ash genome. Therefore, it was aimed to perform genome-wide analysis of *LEA* genes in ash genome using bioinformatic tools in this study.

Materials and methods: Ash and LEA protein sequences were obtained from the Ash Tree Genome and LEAP database respectively. Homologous LEA peptides in ash were found using CLC Genomic Workbench 11. Properties of ash LEA proteins were determined with Expasy PROTPARAM. MEGA7 was used to construct the phylogenetic tree. Functional analysis of ash LEA proteins was carried out via Blast2GO. miRNAs targeting transcripts of ash LEA proteins were detected with psRNATarget. The three-dimensional structures of ash LEA proteins were predicted using PHYRE2.

Main result: 118 *LEA* genes (*FexLEA*) were identified in ash genome. FexLEA were divided into 5 distinct clusters according to phylogenetic analysis. The major molecular function of FexLEA was found as the binding activity. miR838 was the most common miRNA targeting FexLEA transcripts.

Highlights: This study will provide the basis for further functional analysis of LEA proteins in ash.

Keywords: Ash, LEA, Genome Analysis

LEA Genlerinin Dişbudak *(Fraxinus excelsior)* Genomunda Tanımlanması ve Karakterizasyonu

Öz

Çalışmanın amacı: LEA proteinleri, bitkilerin abiyotik stres koşullarına tepki vermesinde önemli rol oynamaktadır. Dişbudak zeytingiller ailesinden genom dizisi tamamlanmış bir ağaç türüdür. Dişbudak genomunda *LEA* genlerine dair açıklamalar bulunsa da kapsamlı bir analiz yoktur. Bu sebeple, bu çalışmada biyoinformatik araçlar kullanılarak *LEA* genlerinin dişbudak genomunda detaylı analizinin yapılması amaçlanmıştır.

Materyal ve yöntem: Ash ve LEA protein dizileri sırasıyla Ash Tree Genome ve LEAP veri tabanından elde edilmiştir. Homolog dişbudak LEA peptitleri, CLC Genomic Workbench 11 kullanılarak bulunmuştur. Dişbudak LEA proteinlerinin özellikleri, Expasy PROTPARAM ile belirlenmiştir. MEGA7 filogenetik ağaç oluşturmak için kullanılmıştır. Dişbudak LEA proteinlerinin fonksiyonel analizi, Blast2GO ile gerçekleştirilmiştir. Dişbudak LEA proteinlerinin transkriptlerini hedef alan miRNA'lar psRNATarget ile tespit edilmiştir. Dişbudak LEA proteinlerinin üç boyutlu yapıları, PHYRE2 kullanılarak tahmin edilmiştir.

Sonuçlar: Dişbudak genomunda 118 tane *LEA* geni (*FexLEA*) tanımlanmıştır. FexLEA'lar filogenetik analize göre 5 dala ayrılmıştır. Fex-LEA üyelerinin başlıca moleküler fonksiyonu bağlanma aktivitesidir. miR838'in FexLEA transkriptlerini hedef alan en yaygın miRNA olduğu görülmüştür.

Önemli vurgular: Bu çalışma, dişbudak LEA proteinlerinin fonksiyonel analizleri için temel sağlayacaktır.

Anahtar Kelimeler: *Dişbudak*, LEA, Genom Analizi

Introduction

Abiotic stress factors induce deleterious changes at the physiological, biochemical and molecular level. Plants have adapted protective mechanisms in response to abiotic stress factors to survive. Accumulation of late embryogenesis abundant (LEA) proteins is one of the stress tolerance mechanisms that protect cells from detrimental effects of abiotic stress factors (Beck, Fettig, Knake, Hartig, & Bhattarai, 2007; Hanin et al., 2011; Sharma & Laxmi, 2016; Verslues, Agarwal, Katiyar-Agarwal, Zhu, & Zhu, 2006). LEA proteins were firstly described in cotton seeds during late developmental stage (Dure III, Greenway & Galau, 1981). Vegetative organs of plants express these proteins in response to stress factors, especially drought. LEA proteins were also found in several organisms including bacteria, fungi, and invertebrates (Battaglia, Olvera-Carrillo, Garciarrubio, Campos, & Covarrubias, 2008). They are divided into eight groups as LEA1-6, dehydrin, dehydrin and SMP (Seed Maturation Protein) according to amino acid sequences and conserved motifs in the PFAM database (Hunault & Jaspard, 2010). Alternative classifications are suggested due to their presence in different organisms and differences in expression profiles. LEA proteins are related to abiotic stress tolerance, however, their exact molecular function is still unclear. *In vitro* experiments suggest a molecular shield role for LEA proteins to prevent protein aggregation. They protect enzymes such as lactate dehydrogenase, citrate synthase, catalase, malate dehydrogenase, fumarase and rhodanese against cold and water stress (Bravo et al., 2003; Goyal, Walton, & Tunnacliffe, 2005; Grelet et al., 2005; Hara, Terashima, & Kuboi, 2001; Honjoh et al., 2000; Reyes et al., 2005). It has been shown that LEA proteins interact with the membrane, to stabilize and protect it from drying (Tolleter et al. 2007). Due to their hydrophilic nature, they are considered as hydration buffers and prevent water loss during dehydration. LEA proteins lack secondary structure in solution, but become folded during dehydration. The diverse functions of LEA proteins are owed to their structural plasticity (Tunnacliffe & Wise, 2007).

Genome-wide identification of LEA proteins with the advancement in plant whole genome sequencing provides evolutionary information and a basis for further structural and functional studies. To date, genome-wide characterization of LEA proteins were performed in several plant species including rice, Arabidopsis, soybean, poplar, sweet orange, potato, tomato, cucumber, maize, Moso bamboo, rapeseed, watermelon/melon (Altunoglu, Baloglu, Yer, Pekol, & Baloglu, 2016; Cao & Li, 2015; Altunoglu, Baloglu, Baloglu, Yer, & Kara, 2017; Charfeddine, Saïdi, Charfeddine, & Gargouri-Bouzid, 2015; Huang et al., 2016; Hundertmark & Hincha, 2008; Lan, Gao, & Zeng, 2013; Li, Xu, Yang, Li, & Hu, 2011; Li & Cao, 2016; Liang et al., 2016; Pedrosa, Martins, Gonçalves, & Costa, 2015; Wang et al., 2007).

Fraxinus excelsior, also known as ash, belongs to olive family Oleaceae. It is dispersed across Europe from the Atlantic coast in the west to Russia in the east, from Norway in the north to Spain, Italy, Greece, and Iran in the south (Semizer-Cuming, Kjær, & Finkeldey, 2017). Even if ash grows more efficiently in moist soils, they are resistant to drought. On the other hand, they are intolerant to cold stress, especially to spring frosts. In this context, identification of LEA proteins in ash gains importance in terms of its drought tolerance and cold intolerance. Whole genome sequence of European ash was reported in 2017 (Sollars et al., 2017). However, there is a lack in the comprehensive analysis of *LEA* genes in ash. In this study, we have performed genome-wide analysis of LEA family in ash and identified 118 LEA proteins in ash genome. Ash LEA proteins were characterized using bioinformatic tools.

Materials and Methods

Identification of LEA Genes in Ash Genome

LEA protein sequences were obtained from the LEAP database which contains information about LEA proteins from plants and other organisms(http://forge.info.univangers.fr/~gh/Leadb/index.php?action=0&mode=0) (Hunault & Jaspard, 2010). Ash protein sequence was retrieved from the Ash Tree Genome database (http://www.ashgenome.org). In order to find

out homologous LEA peptides in ash, LEA protein sequences and ash protein sequence were compared using BLASTP tool of CLC Genomic Workbench 11. In addition, the conserved region analysis of LEA proteins was performed in the PFAM database (http://pfam.sanger.ac.uk/) with the help of CLC Genomic Workbench 11 (Finn et al., 2015). After comparison, the repetitive sequences were removed and possible ash LEA proteins were determined. Properties of ash LEA proteins such as isoelectric point, molecular weight, amino acid length, and stability were determined using the online tool of Expasy PROTPARAM (https://web.expasy.org/protparam/) (Gasteiger et al., 2005).

Sequence Alignment, Phylogenetic Analysis, and İdentification of Conserved Motifs

The amino acid sequences of ash LEA proteins were imported to the MEGA7 software (Kumar, Stecher, & Tamura, 2016) and aligned with the ClustalW algorithm (Larkin et al., 2007). The phylogenetic tree was constructed by Maximum Likelihood Method (MLM) with 1000 bootstrap value. MEME Suite online tool (http://meme-suite.org/) was used to find conserved motifs of ash LEA protein sequences (Bailey & Elkan, 1994).

Gene Ontology Analysis (GO)

Functional analysis of ash LEA proteins was carried out via Blast2GO software (https://www.blast2go.com) (Conesa & Götz, 2008). The output provides information about molecular function, subcellular localization and the biological processes involved.

Orthologous Relationships of Ash LEA Proteins

Amino acid sequences of ash LEA proteins were compared to poplar (*Populus trichocarpa*), eucalyptus (*Eucalyptus grandis*), and sweet orange (*Citrus sinensis*) using BLASTP algorithm in Phytozome 12 database (https://phytozome.jgi.doe.gov) (Goodstein et al., 2011). Hits were chosen according to $\geq E^{-50}$ expectation value and $\geq$ 60% similarity parameters. Duplication analysis of FexLEA proteins was performed using CLC Genomic Workbench 11 with FexLEA protein sequences obtained from Ash Tree Genome database.

Homologous and Nonhomologous Change Rates

Ash LEA proteins and orthologs in poplar, eucalyptus, and sweet orange were aligned using the online tool ClustalOmega (https://www.ebi.ac.uk/Tools/msa/clustalo) (Sievers et al., 2011). Homologous (Ks) and non-homologous (Ka) change rates were calculated through alignment of amino acid sequences of orthologous pairs and their respective cDNA sequences via PAL2NAL online tool (http://www.bork.embl.de/pal2nal/) (Suyama, Torrents, & Bork, P. 2006). Then, using the equation T (million years ago, MYA) = Ks / 2λ ($\lambda = 6.5x10^{-9}$), separation time of orthologous gene pairs in the evolutionary process was calculated (Lynch & Conery, 2000).

MicroRNAs (miRNAs) Targeting LEA Transcripts

All known plant miRNA sequences were obtained from miRBase v21 (http://www.mirbase.org/) (Kozomara & Griffith-Jones, 2013). In order to detect miRNAs targeting transcripts of ash LEA proteins, Plant Small RNA Target Analysis Server, psRNATarget (http://plantgrn.noble.org/psRNATarget) was used (Dai & Zhao, 2011).

Homology Modeling of Ash LEA Proteins

The three-dimensional structures of ash LEA proteins were predicted using Protein Homology/ analogY Recognition Engine V 2.0, PHYRE2 (http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index).

Results and Discussion

Identification of LEA Genes in Ash

LEA protein sequences from LEAP database and ash protein sequence from Ash Tree Genome database were analyzed to identify ash LEA proteins in CLC Genomic Workbench 11. It was found that, the ash has the highest number of LEA proteins (118) identified up to date among plant species (Table S1).

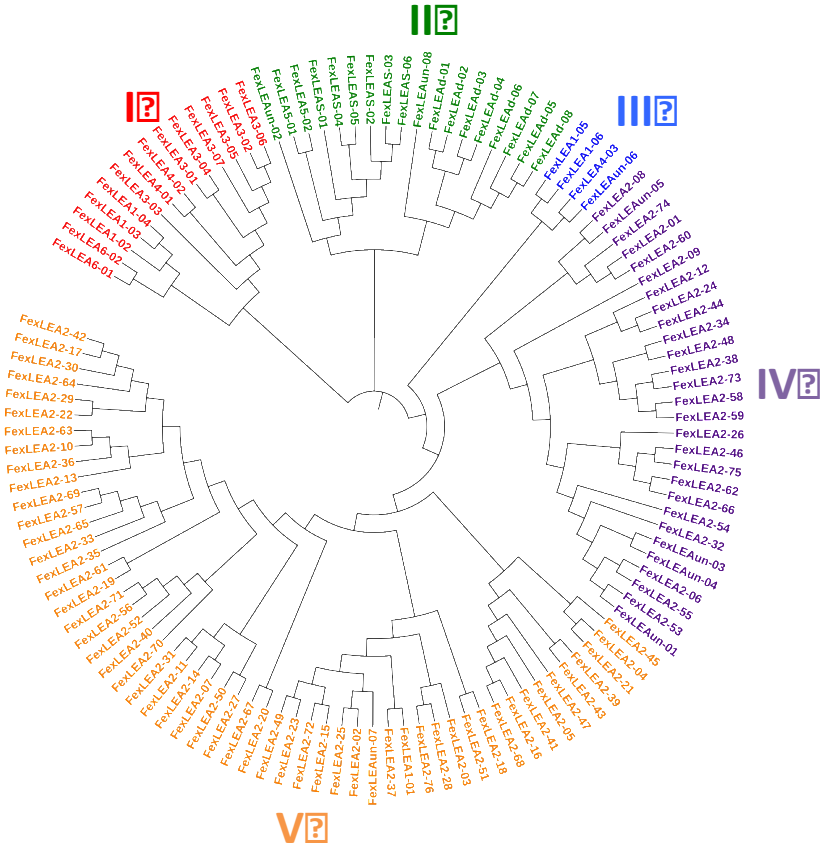

Figure 1. Phylogenetic tree of ash LEA family. The proteins were divided into five distinct clusters and each cluster was differently colored.

These genes were divided into nine groups namely *FexLEA1*, *FexLEA2*, *FexLEA3*, *FexLEA4*, *FexLEA5*, *FexLEA6*, *FexLEAs* (SMP group), *FexLEAd* (dehydrin group) and *FexLEAun* (unknown group). *FexLEA2* was the largest group (76 genes) in ash LEA family. LEA2 group was also reported as the largest group in sweet orange, cucumber, watermelon and melon (Pedrosa et al., 2015; Altunoglu et al., 2016; Altunoglu et al., 2017). For FexLEA proteins; the protein length ranged from 68 amino acids (FexLEA3-01) to 778 amino acids (FexLEA4-01). The molecular weight was between 7.85 kDa (FexLEA3-01) and 85.83 kDa (FexLEA4-01). According to pI values, 74% of the FexLEA proteins were basic. All FexLEA4, FexLEA5, and FexLEA6 were acidic, whereas most of the FexLEA1, FexLEA2, and FexLEA3 were basic. Both the most acidic (pI: 4.64; FexLEA2-60 and the most basic (pI: 10.43; FexLEA2-73) proteins were belong to FexLEA2 group. Similar to our results, LEA proteins were mostly basic in purple false brome, watermelon and melon (Filiz, Ozyigit, Tombuloglu & Koc, 2013; Altunoglu, Baloglu, Baloglu, Yer & Kara, 2017).

Phylogenetic Analysis and İdentification of Conserved Motifs

Phylogenetic analysis was performed to determine the evolutionary relationships of FexLEA proteins. The phylogenetic tree was constructed for 118 FexLEA proteins using the maximum likelihood method with 1000 bootstrap value (Figure 1). According to the phylogenetic tree, ash LEA proteins were divided into 5 main clusters and numbered from Cluster I to Cluster V (Figure 1). The FexLEA2 proteins were distributed to largest

clusters; Cluster IV and Cluster V. All FexLEA1 and FexLEA3 proteins were located in Cluster I. All FexLEA5, FexLEAd and FexLEAs proteins were accumulated in Cluster

Table 1. 30 different motifs identified in FexLEA proteins

Motif No.	Sites	*E*-value	Amino acid sequence composition of the motif	Width
Motif 1	42	4.0e-351	ZVTJTARNPNKKIGIYYDKJSV	22
Motif 2	17	1.2e-400	YASYHGQQITYYTSJPPVYQGHKDVNVWSPFLY GNQVPVAPYNGVALNQD	50
Motif 3	68	7.5e-342	LLIWLVLRPKKPKFSVQSATI	21
Motif 4	17	5.2e-277	TFKVIGRVRWKVGTFVSGRYHLHVNCPAYIPFG N	34
Motif 5	25	5.0e-329	AEVTVKNPNFGSFKYDBSTVTJYYRGDVVGEAV IPAGRAKARSTRKMNIT	50
Motif 6	33	2.4e-157	VFYKDQKLGDGSLPPFYQGRKN	22
Motif 7	5	1.6e-078	MSPILENSPKHCAKQGLNFKKFNKRLFFAFSTFL LTVLSFI	41
Motif 8	7	2.0e-093	FMVQAGSDNTGVPTDMLSMNATVKFTFRNTAT FFGVHVTST	41
Motif 9	9	1.8e-067	RNTGIVIGNAIKYQLSQSCSV	21
Motif 10	11	4.2e-080	GKVELMKVIKKKKSTDMNCTM	21
Motif 11	29	9.7e-067	RRSCCRRCCCF	11
Motif 12	9	2.2e-065	WVPDPVTGYYRPETESKEIDAAELREKLL	29
Motif 13	23	3.3e-056	VPLDLKLKSKVRLKVGAFKTR	21
Motif 14	6	1.4e-081	EDPGAITIGEALEATALSAGNKPVDESDAAAIQA AEVRATGLTHIVP	47
Motif 15	7	3.1e-066	VPLKLEFTIRSKAYVLGKLVKPKFYRKIECSIVID PKKLN	40
Motif 16	4	2.6e-059	AKGAIPFDTTSZISGKFGVFFFDIPLKAKISCEITL NARNQTIDRQSCYP	50
Motif 17	5	9.6e-059	QSAAAYNARTTRDEDKTKLGDVLTDASSKLPSD KPVTREDAEGVIGAEJR	50
Motif 18	4	2.4e-051	VLPFLIAAAYLLWPSDPDLSIVRLRLDGLQFHTR PKISLDI	41
Motif 19	17	2.9e-036	YALNLSSPNJJSSSI	15
Motif 20	8	2.8e-035	RHLKVFCGPLVVPFSSNKRSG	21
Motif 21	10	7.3e-039	KSEEKKGFLEKIKEKLPGGHK	21
Motif 22	6	6.9e-036	RPVYYVQSPSRDSHDGEKTSSSF	23
Motif 23	11	7.3e-030	MADRVLSNPNLYSDJASGTLPLSSYTKJ	28
Motif No.	Sites	*E*-value	Amino acid sequence composition of the motif	Width
Motif 24	4	4.0e-028	JKANFAAAGSYVDKWVVDGINWDRSKNGN	29
Motif 25	4	2.8e-027	RRKGEKGWKECDVIEEEGLLDDEDGDKGFPRRC YLLAFVVGFFVLFSFFA	50
Motif 26	7	6.2e-027	EKEZEEKKHEELLEKLHRSGSSSSSSSDE	29
Motif 27	4	2.5e-026	NPYPVPIPICDINYSLKSAGRVIASGKIPDPGSLKG NDNTMLDVPIKVPH	50
Motif 28	6	3.1e-021	PSYQNYPYNQQSYZHDPDSAR	21
Motif 29	4	8.5e-024	YLDLNLAARDVSLYGAVARDLKLEKYSGN	29
Motif 30	4	1.5e-020	SPDGEVKEKKGFLDKIKEKJPGYHSKTEE	29

II. FexLEA1 proteins were observed in Cluster III and Cluster V, while FexLEA4 proteins were found in Cluster I and Cluster III. Presence of different motifs may cause distribution of FexLEA1, FexLEA2 and FexLEA4 proteins to different branches in the phylogenetic tree. Therefore, amino acid compositions of motifs found in ash LEA proteins were analyzed by MEME software. 30 different motifs were identified for 105.

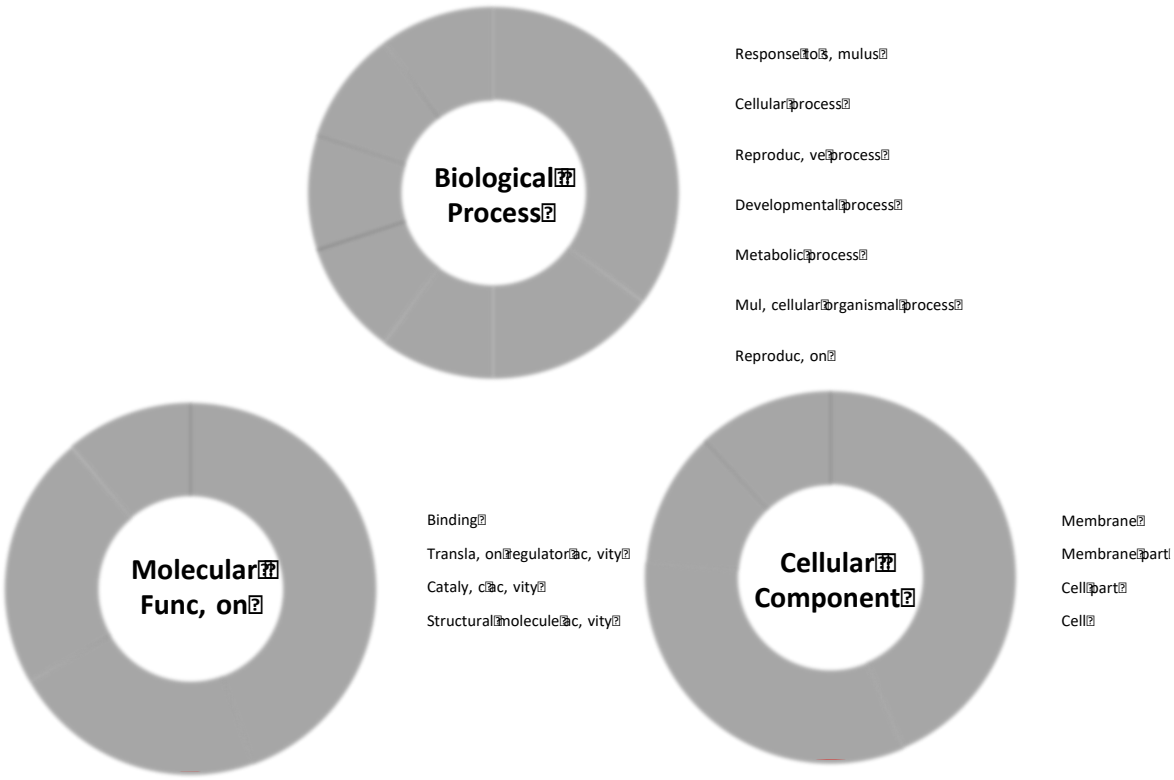

Figure 2. Gene ontology analysis for FexLEA proteins

FexLEA proteins (Table 1; Figure S1). All FexLEA3 proteins shared motif 12, all FexLEAd proteins shared motif 21 and motif 26, and all FexLEAs group proteins shared motif 14 and motif 17. Diverse motifs were detected in FexLEA1 and FexLEA2 proteins as expected. LEA4, LEA5 and LEA6 group proteins had no motifs. It is noteworthy that the proteins that had a similar motif organization were in the same cluster in the phylogenetic tree. For instance, FexLEA2-57, FexLEA2-65, FexLEA2-69, FexLEA2-33, and FexLEA35 had the same motif organization and they are located in the same sub-group of Cluster V.

Gene Ontology Analysis

Functional analysis of FexLEA proteins was carried out via Blast2GO software. After FexLEA amino acid sequences were uploaded, their molecular function, cellular localization, and the biological role were determined (Figure 3, Table S2). As a result of gene ontology analysis, FexLEA proteins were found to play a role in 7 different biological processes including response to a stimulus, cellular process, reproductive process, developmental process, metabolic process, multicellular organismal process, and reproduction. Half of the proteins were involved in response to stimuli and cellular processes which supports their role in the stress tolerance mechanism (Beck et al., 2007; Hanin et al., 2011; Sharma & Laxmi, 2016; Verslues et al., 2006). Most of the FexLEA proteins were localized to membrane and membrane parts. Other FexLEA proteins were distributed to cell and cell parts. The major molecular function for FexLEA proteins was found as the binding activity, which is in line with the binding property of this protein family to Ca^{+2} ions, and metal ions (Alsheikh,

Svensson, & Randall, 2005; Krüger, Berkowitz, Stephan, & Hell, 2002).

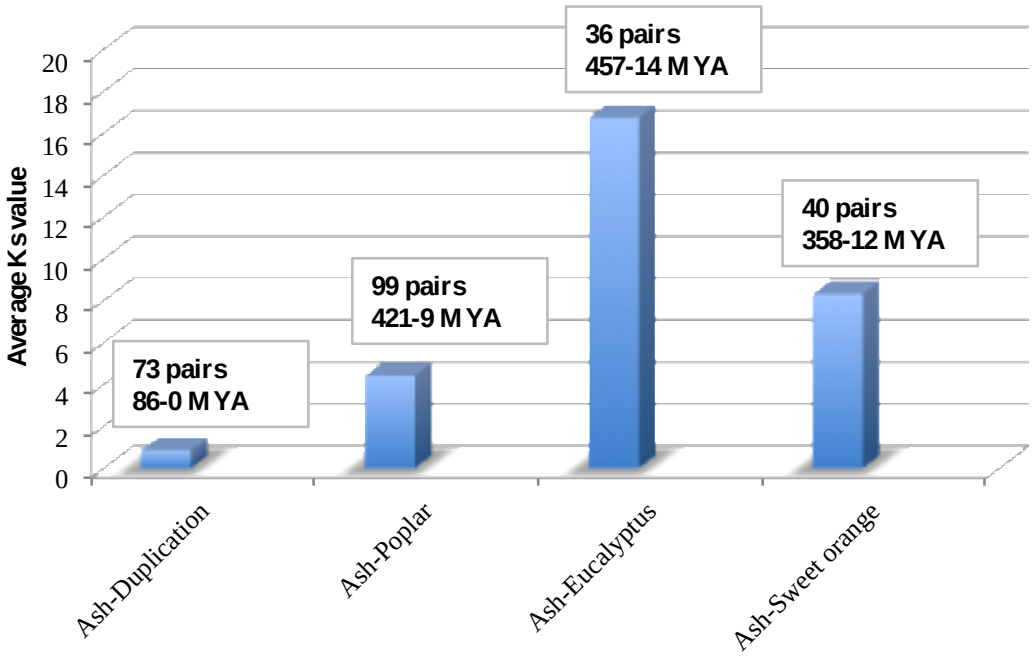

Figure 3. Estimation of duplication and divergence times of FexLEA genes with orthologous *LEA* gene pairs between ash and poplar, eucalyptus and sweet orange

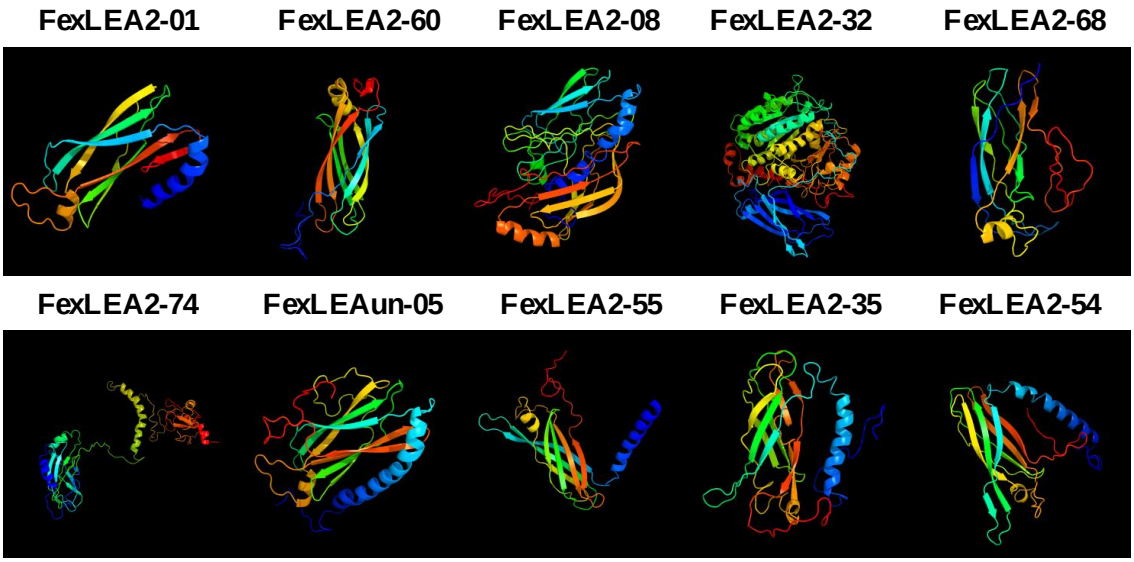

Figure 4. Predicted strctures of FexLEA proteins

Besides, these proteins interact with phospholipid bilayer of membranes to protect them from drying (Tolleter et al., 2007). They have also catalytic activity, translational regulator activity, and structural molecule function activity.

Duplication Events and Divergence Rates of FexLEA Genes

73 pairs of FexLEA genes were duplicated with divergence time varied between 86-0 MYA. In addition, a total of 99, 36 and 40 pairs of orthologous LEA genes were

identified between ash and eudicot plants poplar, eucalyptus, sweet orange respectively (Figure 3, Table S3). The earliest divergence time was observed between ash and eucalyptus with an average of 128 MYA and the latest divergence time was observed between ash and poplar with an average of 34 MYA (Table S3). These results have shown that poplar and ash are closer genetically than other analyzed plants.

miRNAs Targeting FexLEA *Transcripts*

miRNAs are small RNA molecules regulating gene expression. They are evolutionarily conserved in plant species and play a major role in abiotic stress response mechanisms (Zhang, 2015). Therefore, miRNAs targeting FexLEA transcripts were identified by psRNA Target Server (http://plantgrn.noble.org/psRNATarget/). We have found that, 102 different FexLEA transcripts were targeted by 185 different Arabidopsis miRNAs (Table S4). miR838, miR865-3p, miR4221, miR5658 and miR414 were the most common miRNAs targeting FexLEA transcripts. FexLEA2-75 was the most targeted transcript by different miRNAs (targeted by 16 different miRNAs). miR838 is a conserved miRNA found in ginger, *Brassica juncea*, Arabidopsis, *Salvia miltiorrhiza* and *Brassica napus* (Ramachandran & Chen, 2008; Shao, Qiu, & Lu, 2015; Singh, Srivastava, & Sharma, 2016; Srivastava, Srivastava, Suprasanna, & D'souza, 2012; Wang, Qiao, Zhang, Shi, & Zhang, 2017). In wheat, miR838 was shown to regulate drought response by targeting heat shock protein gene (Akdogan, Tufekci, Uranbey, & Unver, 2016). miR5658 has been observed in ginger, blueberry, *Ferula gumosa* and regulate drought tolerance in tomato (Candar-Cakir, Arican, & Zhang, 2016; Li et al., 2014; Singh et al., 2016; Najafabadi & Naghavi, 2018). miR414, another conserved miRNA among plant species, was associated with drought resistance in *Physcomitrella patens* (Wan et al., 2011). All these miRNAs might reprogramme FexLEA gene expression and regulate their effect in stress tolerance mechanisms.

3D-structures of FexLEA Proteins

The 3D-structures of the FexLEA proteins were estimated based on > 90% confidence interval and > 70% similarity parameters. 10 FexLEA proteins (FexLEA2-01, FexLEA2-60, FexLEA2-08, FexLEA2-32, FexLEA2-68, FexLEA2-74, FexLEAun-05, FexLEA2-55, FexLEA2-34, FexLEA2-54) were modeled with these criteria. Although β -sheet structures were predominant, α-helices were also detected in FexLEA proteins (Figure 4). In accordance with the present result, NMR results have shown that Arabidopsis LEA14 protein has αβ-fold consisting of one α-helix and seven β-strands (Singh et al., 2005). β-sheet dominant secondary structure was also reported in cucumber, watermelon, and melon LEA proteins (Altunoglu et al., 2016; Altunoglu et al., 2017).

Conclusions

In the present study, a total of 118 FexLEA encoding genes were identified and characterized by phylogenetic analysis, motif search, miRNA identification, duplication events and estimation of 3D structure. The results have demonstrated that LEA constitutes a large and diverse family of proteins in ash. These results will provide comprehensive insight for future functional analysis of LEA proteins to unravel the role of LEA proteins in stress tolerance mechanisms.

References

Akdogan, G., Tufekci, E. D., Uranbey, S. & Unver, T. (2016). miRNA-based drought regulation in wheat. *Functional & Integrative Genomics*, 16(3), 221-233.

Alsheikh, M. K., Svensson, J. T. & Randall, S. K. (2005). Phosphorylation regulated ion-binding is a property shared by the acidic subclass dehydrins. *Plant, Cell & Environment*, 28(9), 1114-1122.

Altunoglu, Y. C., Baloglu, M. C., Baloglu, P., Yer, E. N. & Kara, S. (2017). Genome-wide identification and comparative expression analysis of LEA genes in watermelon and melon genomes. *Physiology and Molecular Biology of Plants*, 23(1), 5-21.

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 Regulation*, 80(2), 225-241.

Bailey, T. L. & Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in bipolymers. *Proceedings of the 2nd International Conference on Intelligent Systems for Molecular Biology*, 148, 28-36.

Battaglia, M., Olvera-Carrillo, Y., Garciarrubio, A., Campos, F. & Covarrubias, A. A. (2008). The enigmatic LEA proteins and other hydrophilins. *PlantPphysiology*, 148(1), 6-24.

Beck, E. H., Fettig, S., Knake, C., Hartig, K. & Bhattarai, T. (2007). Specific and unspecific responses of plants to cold and drought stress. *Journal of Biosciences*, 32(3), 501-510.

Bravo, L. A., Gallardo, J., Navarrete, A., Olave, N., Martínez, J., Alberdi, M., ... & Corcuera, L. J. (2003). Cryoprotective activity of a cold-induced dehydrin purified from barley. *Physiologia Plantarum*, 118(2), 262-269.

Candar- Cakir, B., Arican, E. & Zhang, B. (2016). Small RNA and degradome deep sequencing reveals drought- and tissue- specific micrornas and their important roles in drought- sensitive and drought- tolerant tomato genotypes. *Plant Biotechnology Journal*, 14(8), 1727-1746.

Cao, J. & Li, X. (2015). Identification and phylogenetic analysis of late embryogenesis abundant proteins family in tomato (Solanum lycopersicum). *Planta*, 241(3), 757-772.

Charfeddine, S., Saïdi, 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. *Molecular Biology Reports*, 42(7), 1163-1174.

Conesa, A. & Götz, S. (2008). Blast2GO: A comprehensive suite for functional analysis in plant genomics. *International Journal of Plant Genomics*, 2008(2008,619832), 1-12.

Dai, X. & Zhao, P. X. (2011). psRNATarget: a plant small RNA target analysis server. *Nucleic Acids Research*, *39*(suppl_2), W155-W159.

Dure III, L., 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(14), 4162-4168.

Filiz, E., Ozyigit, I. I., Tombuloglu, H. & Koc, I. (2013). In silico comparative analysis of LEA (Late Embryogenesis Abundant) proteins in Brachypodium distachyon L. *Plant Omics*, 6(6), 433.

Finn, R. D., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L. & Salazar, G. A. (2015). The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Research*, 44(D1), D279-D285.

Gasteiger, E., Hoogland, C., Gattiker, A., Wilkins, M. R., Appel, R. D. & Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server. *The proteomics protocols handbook* (pp. 571-607). Humana press.

Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J. & Rokhsar, D. S. (2011). Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Research*, 40(D1), D1178-D1186.

Goyal, K., Walton, L. J. & Tunnacliffe, A. (2005). LEA proteins prevent protein aggregation due to water stress. *Biochemical Journal*, 388(1), 151-157.

Grelet, J., Benamar, A., Teyssier, E., Avelange-Macherel, M. H., Grunwald, D. & Macherel, D. (2005). Identification in pea seed mitochondria of a late-embryogenesis abundant protein able to protect enzymes from drying. *Plant Physiology*, 137(1), 157-167.

Hanin, M., Brini, F., Ebel, C., Toda, Y., Takeda, S. & Masmoudi, K. (2011). Plant dehydrins and stress tolerance: versatile proteins for complex mechanisms. *Plant Signaling & Behavior*, 6(10), 1503-1509.

Hara, M., Terashima, S. & Kuboi, T. (2001). Characterization and cryoprotective activity of cold-responsive dehydrin from Citrus unshiu. *Journal of Plant Physiology*, 158(10), 1333-1339.

Honjoh, K. I., Matsumoto, H., Shimizu, H., Ooyama, K., Tanaka, K., Oda, Y., ... & Iio, M. (2000). Cryoprotective activities of group 3 late embryogenesis abundant proteins from Chlorella vulgaris C-27. *Bioscience, Biotechnology, and Biochemistry*, 64(8), 1656-1663.

Huang, Z., Zhong, X. J., He, J., Jin, S. H., Guo, H. D., Yu, X. F. & 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(11), e0165953.

Hunault, G. & Jaspard, E. (2010). LEAPdb: a database for the late embryogenesis abundant proteins. *Bmc Genomics*, 11(1), 221.

Hundertmark, M. & Hincha, D. K. (2008). LEA (late embryogenesis abundant) proteins and their encoding genes in Arabidopsis thaliana. *BMC Genomics*, 9(1), 118.

Kozomara, A. & Griffiths-Jones, S. (2013). miRBase: annotating high confidence microRNAs using deep sequencing

data. *Nucleic Acids Research*, 42(D1), D68-D73.

Krüger, C., Berkowitz, O., Stephan, U. W. & Hell, R. (2002). A metal-binding member of the late embryogenesis abundant protein family transports iron in the phloem of Ricinus communis L. *Journal of Biological Chemistry*, 277(28), 25062-25069.

Kumar, S., Stecher, G. & Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, 33(7), 1870-1874.

Lan, T., Gao, J. & Zeng, Q. Y. (2013). Genome-wide analysis of the LEA (late embryogenesis abundant) protein gene family in Populus trichocarpa. *Tree Genetics & Genomes*, 9(1), 253-264.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H. & Thompson, J. D. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics*, 23(21), 2947-2948.

Li, X. & Cao, J. (2016). Late Embryogenesis Abundant (LEA) gene family in maize: identification, evolution, and expression profiles. *Plant Molecular Biology Reporter*, 34(1), 15-28.

Li, X., Hou, Y., Zhang, L., Zhang, W., Quan, C., Cui, Y. & Bian, S. (2014). Computational identification of conserved microRNAs and their targets from expression sequence tags of blueberry (Vaccinium corybosum). *Plant Signaling & Behavior*, 9(9), e29462.

Li, L., Xu, H., Yang, X., Li, Y. & Hu, Y. (2011). Genome-wide identification, classification and expression analysis of LEA gene family in soybean. Scientia Agricultura Sinica, 44(19), 3945-3954.

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

Lynch, M. & Conery, J. S. (2000). The evolutionary fate and consequences of duplicate genes. *Science*, 290(5494), 1151-1155.

Najafabadi, A. S. & Naghavi, M. R. (2018). Mining Ferula gummosa transcriptome to identify miRNAs involved in the regulation and biosynthesis of terpenes. *Gene*, 645, 41-47.

Pedrosa, A. M., Martins, C. D. P. S., Gonçalves, L. P. & Costa, M. G. C. (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(12), e0145785.

Ramachandran, V., & Chen, X. (2008). Small RNA metabolism in Arabidopsis. *Trends in Plant Science*, 13(7), 368-374.

Reyes, J. L., Rodrigo, M. J., Colmenero-Flores, J. M., Gil, J. V., Garay-Arroyo, A., Campos, F., ... & Covarrubias, A. A. (2005). Hydrophilins from distant organisms can protect enzymatic activities from water limitation effects in vitro. *Plant, Cell & Environment*, 28(6), 709-718.

Semizer-Cuming, D., Kjær, E. D. & Finkeldey, R. (2017). Gene flow of common ash (Fraxinus excelsior L.) in a fragmented landscape. *PloS One*, 12(10), e0186757.

Shao, F., Qiu, D. & Lu, S. (2015). Comparative analysis of the Dicer-like gene family reveals loss of miR162 target site in SmDCL1 from Salvia miltiorrhiza. *Scientific Reports*, 5, 9891.

Sharma, M. & Laxmi, A. (2016). Jasmonates: emerging players in controlling temperature stress tolerance. *Frontiers in Plant Science*, 6, 1129.

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W. & Thompson, J. D. (2011). Fast, scalable generation of high- quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, 7(1), 539.

Singh, N., Srivastava, S. & Sharma, A. (2016). Identification and analysis of miRNAs and their targets in ginger using bioinformatics approach. *Gene*, 575(2), 570-576.

Singh, S., Cornilescu, C. C., Tyler, R. C., Cornilescu, G., Tonelli, M., Lee, M. S. & Markley, J. L. (2005). Solution structure of a late embryogenesis abundant protein (LEA14) from Arabidopsis thaliana, a cellular stress-related protein. *Protein Science*, 14(10), 2601-2609.

Sollars, E. S., Harper, A. L., Kelly, L. J., Sambles, C. M., Ramirez-Gonzalez, R. H., Swarbreck, D., ... & Worswick, G. (2017). Genome sequence and genetic diversity of European ash trees. *Nature*, 541(7636), 212.

Srivastava, S., Srivastava, A. K., Suprasanna, P. & D'souza, S. F. (2012). Identification and profiling of arsenic stress-induced microRNAs in Brassica juncea. *Journal of Experimental Botany*, 64(1), 303-315.

Suyama, M., Torrents, D. & Bork, P. (2006). PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Research*, 34(suppl_2), W609-W612.

Tolleter, D., Jaquinod, M., Mangavel, C., Passirani, C., Saulnier, P., Manon, S. & Macherel, D. (2007). Structure and function of a mitochondrial late embryogenesis abundant protein are revealed by desiccation. *The Plant Cell*, 19(5), 1580-1589.

Tunnacliffe, A. & Wise, M. J. (2007). The continuing conundrum of the LEA proteins. *Naturwissenschaften*, 94(10), 791-812.

Verslues, P. E., Agarwal, M., Katiyar- Agarwal, S., Zhu, J. & Zhu, J. K. (2006). Methods and concepts in quantifying resistance to drought, salt and freezing, abiotic stresses that affect plant water status. *The Plant Journal*, 45(4), 523-539.

Wan, P., Wu, J., Zhou, Y., Xiao, J., Feng, J., Zhao, W. & Chen, J. Y. (2011). Computational analysis of drought stress-associated miRNAs and miRNA co-regulation network in Physcomitrella patens. *Genomics, Proteomics & Bioinformatics*, 9(1), 37-44.

Wang, Z., Qiao, Y., Zhang, J., Shi, W. & Zhang, J. (2017). Genome wide identification of microRNAs involved in fatty acid and lipid metabolism of Brassica napus by small RNA and degradome sequencing. *Gene*, 619, 61-70.

Wang, X. S., Zhu, H. B., Jin, G. L., Liu, H. L., Wu, W. R. & Zhu, J. (2007). Genome-scale identification and analysis of LEA genes in rice (Oryza sativa L.). *Plant Science*, 172(2), 414-420.

Zhang, B. (2015). MicroRNA: a new target for improving plant tolerance to abiotic stress. *Journal of Experimental Botany*, 66(7), 1749-1761.