

沙冬青 *AmLEA5* 基因的生物信息学分析及非生物胁迫下的表达模式

赵晓鑫, 刘美芹, 师 静, 智冠华, 刘胜利, 张俊琦, 陈玉珍, 卢存福

(北京林业大学生物科学与技术学院/林木育种国家工程实验室/教育部林木花卉遗传育种重点实验室, 北京 100083)

摘要: 对用固相扣除杂交方法从低温驯化沙冬青克隆得到的 *AmLEA5* 基因进行表达模式分析和生物信息学分析, 表明该基因编码一种第 5 族胚胎发育晚期丰富蛋白 (LEA), 全长 693 bp, 含有 1 个 297 bp 的开放阅读框, 编码 98 个氨基酸, 预测 *AmLEA5* 的分子量为 10.6 kDa, 是一种亲水性蛋白, 有多个磷酸化位点。密码子偏好性分析表明该基因略偏好于用 A 或 T 结尾的密码子。系统发生分析表明, *AmLEA5* 蛋白与菘蓿苜蓿 LEA (ACJ84182.1) 亲缘关系最近。qRT-PCR 结果显示 *AmLEA5* 的表达量在低温、干旱、盐和热胁迫条件下均有上调, 尤其在低温胁迫后期富集量最高。亚细胞定位表明, 用 YFP 标记的 *AmLEA5* 位于细胞质和细胞核内。一系列试验结果表明 *AmLEA5* 基因在沙冬青抵御非生物胁迫, 尤其是在抵御低温胁迫机制中发挥重要作用。

关键词: 沙冬青; *AmLEA5*; 生物信息学分析; 表达模式; 亚细胞定位

Bioinformatic Analysis and Expression Patterns under Abiotic Stresses of *AmLEA5* from *Ammopiptanthus mongolicus*

ZHAO Xiao-xin, LIU Mei-qin, SHI Jing, ZHI Guan-hua, LIU Sheng-li, ZHANG Jun-qi, CHEN Yu-zhen, LU Cun-fu

(College of Biological Sciences and Technology, National Engineering Laboratory for Tree Breeding,

Key Laboratory of Genetics and Breeding in Forest Tree and Ornamental Plants of

Education Ministry, Beijing Forestry University, Beijing 100083)

Abstract: Molecular characterization and expression patterns of *AmLEA5* (AAW31666.1) under abiotic stresses which cloned from cold-and drought-stressed *Ammopiptanthus mongolicus* by a modified solid subtractive hybridization method were reported in this study. Bioinformatic analysis showed that the gene encoded a group 5 late embryogenesis abundant protein (LEA). The overall length of *AmLEA5* was 693 bp, and the cDNA contained a 297 bp ORF encoding a polypeptide of 98 amino acids with a calculated molecular mass of 10.6 kDa. *AmLEA5*, with more phosphorylation sites, was a kind of hydrophilic protein. The analysis of codon bias showed that *AmLEA5* gene was bias toward the synonymous codons with A and T at the third codon position. Phylogenetic analysis indicated that *AmLEA5* protein was most closed to *Medicago truncatula* LEA (ACJ84182.1) in genetic relationship. qRT-PCR analysis demonstrated that *AmLEA5* was up-regulated by cold, drought, salt and hot stresses. Specifically, transcription of *AmLEA5* accumulated more at the late stage of cold treatment. Intracellular localization analysis revealed that *AmLEA5* proteins tagged with yellow fluorescent proteins (YFP) was more likely localized in the cytoplasm and nucleus. The results suggested that *AmLEA5* may played important roles in abiotic stresses responses, especially in cold tolerant mechanism in *A. Mongolicus*.

Key words: *Ammopiptanthus mongolicus*, *AmLEA5*, bioinformatic analysis, expression pattern, subcellular localization

收稿日期: 2012-10-26 修回日期: 2012-11-19 网络出版日期: 2013-04-02

URL: <http://www.cnki.net/kcms/detail/11.4996.S.20130402.1733.002.html>

基金项目: 国家自然科学基金(31270737, 30671476); 北京市自然科学基金(6112016)

作者简介: 赵晓鑫, 硕士研究生, 研究方向: 植物细胞分子生物学。E-mail: xiaoxin115987@sina.com

通信作者: 卢存福, 博士, 教授, 研究方向: 植物细胞分子生物学。E-mail: lucunfu@bjfu.edu.cn

胚胎发育晚期丰富蛋白(LEA, Late embryogenesis abundant proteins)最早发现于棉花子叶中^[1],随后在小麦^[2]、大豆^[3]、玉米^[4]等作物种子中陆续被发现。LEA 蛋白在被子植物、裸子植物,以及苔藓和蕨类植物中都存在,此外在酵母、线虫等物种中也表达^[5]。在植物中,它伴随着种子的成熟过程而产生,大量积累于种子成熟后期特定阶段,在种子萌发后很快消失。一般认为,LEA 蛋白与种子抵抗脱水有关,主要参与维持种子的生长活力^[6],此外,LEA 蛋白的积累参与植物非生物胁迫防御系统^[7],当植物遭受外界逆境时会在细胞中积累大量该类保护蛋白来抵抗胁迫。研究发现低温、干旱、紫外线及机械损伤均可诱导 LEA 蛋白的表达^[8-9]。

根据氨基酸序列或亲水性可以将 LEA 蛋白分成 6 个家族^[10],其中第 5 族与第 3 族在结构和功能上较相似,具有中和离子的功能,但是第 5 族 LEA 缺乏高度的残基专一性,没有明显的保守结构域。组成第 5 族 LEA 蛋白的大多数氨基酸为碱性、亲水性氨基酸,没有半胱氨酸和色氨酸。第 5 族 LEA 序列可形成大量 α -螺旋结构^[11],提供一个亲水表面。该基元序列另外一个显著特点是,无规卷曲结构的含量也比较丰富^[12]。第 5 族 LEA 蛋白之间同源性较低,热稳定性差,沸水浴后该族 LEA 蛋白不能溶于水,一般在种子中含量较高,而且在胁迫条件下会诱导表达,N. Dorit 等^[13]发现柑橘在干旱、热和盐胁迫后会诱导 *LEA5* 基因的表达。

沙冬青(*Ammopiptanthus mongolicus* (Maxim.) Cheng f.)又叫蒙古沙冬青、蒙古黄花木,是古老的第 3 纪残遗物种,可作为观赏植物和药物资源植物^[14-16]。主要分布于内蒙古西部沙漠,是亚洲中部荒漠地区特有的常绿阔叶灌木,具有极强的抗逆境能力,是一种研究林木抗逆机理的理想材料^[17-20]。本实验室前期用优化的固相扣除杂交策略从低温驯化后、抗冻性提高的沙冬青幼苗中克隆了胚胎晚期发生丰富蛋白基因 *AmLEA5*^[21],本研究对该基因的分子生物学特性进行分析。用实时荧光定量 PCR 方法检测 *AmLEA5* 基因在遭受低温、干旱、盐、热胁迫下表达模式的变化;构建表达载体并转化洋葱表皮细胞,检测了 YFP-*AmLEA5* 融合蛋白的亚细胞表达定位,期望为理解木本植物的抗逆性机理提供新的资料,为植物抗逆育种提供新的基因资源。

1 材料与方法

1.1 *AmLEA5* 基因的生物信息学分析

通过优化的固相扣除杂交策略从低温驯化后、抗冻性提高的沙冬青幼苗中克隆到了 *NIOH6* 基因^[21],经序列分析后命名为 *AmLEA5*。用 NCBI 的 BLAST 和 ORF Finder 在线软件分析该序列的同源性及开放阅读框。用 ProtParam 软件分析该基因所编码蛋白的分子量和等电点。用 Signal P 软件预测该蛋白序列中的信号肽的剪切位点。用 TMpred 软件预测该蛋白跨膜区域及膜上的走向。用 SOPMA 和 Predictprotein 软件对该蛋白二级结构及其亲水性进行预测。用 SbuLocV 1.0 和 Target P 软件预测该蛋白的亚细胞定位。用 Clustal W 2 软件构建系统进化树。

密码子偏好性分析使用 EMBOSS 在线程序(<http://vm-bioinfo.toulouse.inra.fr/emboss>)中的 CHIPS (Codon heterozygosity in a protein coding sequence)和 CUSP (Create a codon usage table)以及 Codon W 在线程序(<http://mobyle.pasteur.fr/cgi-bin/portal.py?form=codonw>)。运用软件 CHIPS 对沙冬青 *AmLEA5* 基因进行有效密码子数(Enc)统计,然后运用 CUSP 和 Codon W 计算密码子中 CDS 区的 GC 含量和第 3 位碱基的 GC 含量(GC、GC3s)、密码子使用频率(frequency)和相对同义密码子使用度(RSCU,relatively synonymous codon usage)^[22-23]。

1.2 沙冬青幼苗培养及胁迫处理

用自来水浇透珍珠岩基质,将沙冬青种子埋入该基质中,于恒温箱中 25 °C、光照 16 h、70%湿度培养,中间过程不再浇水。培养 10 d,挑选长势一致的沙冬青幼苗,用自来水浇透,1 h 后取样作为对照,其余的幼苗进行以下处理。所有处理均取地上部分作为试验材料。低温处理:将沙冬青幼苗放入 4 °C 黑暗培养室中,分别在第 1、2、4、8、16 天取样^[24]。干旱处理:将沙冬青幼苗根部珍珠岩小心去掉,放入 15 cm × 15 cm 培养皿中,黑暗条件下,分别在第 1、2、3、4、5 天取样。盐处理:将沙冬青幼苗置于 25 °C、16 h 光照、70%湿度条件下培养,每 7 天浇 1 次 1%的 NaCl 溶液,每次浇 200 mL,分别在第 1、2、4、8、16 天取样。热处理:将沙冬青幼苗置于 40 °C、16 h 光照、70%湿度条件下培养,分别在第 1、2、4、8、12、24 h 取样。以上样品用液氮速冻,然后保存于 -80 °C 冰箱中。

1.3 沙冬青总 RNA 提取

使用 Plant RNA assistant Kit(北京科百奥)试剂盒,按照操作说明进行总 RNA 的提取。从 $-80\text{ }^{\circ}\text{C}$ 冰箱中取出样品,液氮充分研磨后将粉末收集到 1.5 mL 的离心管中,加入 750 μL RNA assistant Lysis Buffer 及 500 μL RNA assistant Extraction Buffer,用振荡器剧烈震荡 5 min 充分混匀,12000 r/min 离心 5 min。取上清加入 500 μL RNA assistant Extraction Buffer,振荡器剧烈震荡 5 min 充分混匀,12000 r/min 离心 5 min。去上清加入等体积氯仿,剧烈震荡 10 min 后 12000 r/min 离心 10 min。取上清加入 1/2 体积的 500 μL RNA assistant Precipitation Buffer,混匀后置于冰上沉淀 20 min。12000 r/min 离心 10 min,弃上清。用 1 mL 80% 乙醇覆盖沉淀,13000 r/min 离心 5 min,弃上清,重复 1 次。挥发乙醇 4 min,加入适量 RNase-free ddH₂O 溶解沉淀。最后用 Nanodrop ND-1000 spectrophotometer (NanoDrop. USA)检测 RNA 浓度。

1.4 cDNA 合成

使用 Invitrogen 公司的 SuperScriptTM III Reverse Transcriptase 试剂盒进行 cDNA 的合成。将 1 μL Oligo(dT)₂₀ (50 $\mu\text{mol/L}$) (USA, Romega), 5 μg 总 RNA, 1 μL 的 10 mmol/L dNTP mix (Takara, Japan) 混匀后, 65 $^{\circ}\text{C}$ 孵育 5 min, 然后立即放到冰上速冷至少 1 min; 在以上的混合物中加入 1 μL RNase Inhibitor (Takara, Japan), 4 μL 的 5 Firststand Buffer, 1 μL 的 0.1 mol/L DTT, 1 μL SuperScriptTM III RT (200 units/ μL), 吸打混匀, 50 $^{\circ}\text{C}$ 孵育 60 min, 70 $^{\circ}\text{C}$ 孵育 15 min。最后将合成好的 cDNA 放在 $-20\text{ }^{\circ}\text{C}$ 保存。

1.5 实时荧光定量 PCR

实时荧光定量 PCR, 即 qRT-PCR, 以 *eIF1* 和 *eIF3* 作为内参基因^[25]。qRT-PCR 引物严格按照荧光定量 PCR 引物设计要求, 使用 Primer 3 软件, 根据 *AmLEA5*、*eIF1* 和 *eIF3* 基因序列进行引物设计, 然后由北京奥科公司合成引物。引物序列分别为:

eIF1F CTGACATGCCCGTAGGAACG; eIF1R CCC-TGCTTATGCCAGTCTTTT

eIF3F TTACATCGTCATCAAGGCTCG; eIF3R ACA-ATTATGGGAAGACGGCAC

R-LEAU TGAATCACTAATGGCTCGCTCCTTC;
R-LEAD ACACCTTGTGCGTTGTGGCCCTC

qRT-PCR 反应所用的仪器是 ABI PRISM7900HT Real-Time PCR System, 使用美国 ABI 公司生产的 384 孔板和贴膜。荧光定量试剂盒是 SYBR Premix

Ex TaqTM kit (Takara, Japan)。20 μL 反应体系包括: 10 μL SYBR Premix Ex Taq, 7.2 μL PCR-grade water, 2 μL cDNA (50ng/ μL), 0.4 μL 的 10 $\mu\text{mol/L}$ 的正向引物和反向引物。荧光定量循环体系参数设为 95 $^{\circ}\text{C}$ 30 s 循环 1 次; 95 $^{\circ}\text{C}$ 变性 5 s, 60 $^{\circ}\text{C}$ 延伸 34 s, 循环 40 次; 然后进行溶解曲线测定, 95 $^{\circ}\text{C}$ 15 s, 60 $^{\circ}\text{C}$ 1 min, 95 $^{\circ}\text{C}$ 15 s, 循环 1 次。每个样品重复 3 次。

基因的表达水平通过 Ct (cycle threshold) 值来呈现。Ct 值是指每个反应管内的荧光信号达到设定的值域时所经历的循环数。反应结束后, 进行溶解曲线分析。用 *eIF1* 和 *eIF3* 基因 Ct 值的几何平均数作为内参, 采用 $2^{-\Delta\text{Ct}}$ 算法计算 *AmLEA5* 基因的相对表达量。根据 Excel 表格中 T-test 对表达数值进行分析, 当 $P < 0.05$ 时, 说明表达出现了明显变化^[26]。

1.6 AmLEA5 基因的亚细胞定位

用 GatewayTM 重组系统^[27] 构建表达载体, 使得融合蛋白 YFP-*AmLEA5* 能够高效而精确表达。用基因枪转化洋葱内表皮, 使其瞬时表达。在激光共聚焦显微镜 (Leica-SP5) 下观察 YFP-*AmLEA5* 融合蛋白在洋葱内表皮细胞中的定位。

2 结果与分析

2.1 AmLEA5 生物信息学分析

通过优化的固相扣除杂交技术得到沙冬青低温诱导基因 *AmLEA5*, 生物信息学分析发现 *AmLEA5* 的 cDNA 全长为 693 bp, 经 ORF Finder 程序分析, 该序列含有 1 个 297 bp 的开放阅读框, 编码 98 个氨基酸 (图 1)。该基因 5' 端非编码区 (UTR, untranslated region) 长 119 bp, 3' UTR 长 277 bp, 其中包括一段 poly (A) 尾。用 ProtParam 预测显示 *AmLEA5* 基因所编码的蛋白分子量为 10.6 kDa, 等电点为 10.19, 由 98 个氨基酸残基组成, 大多数为碱性、亲水性氨基酸, 富含 Ala (11.22%)、Lys (10.20%)、Ser (11.22%)、Val (9.18%), 无 Cys 和 Trp。该基因与多个第 5 族胚胎发育晚期丰富蛋白 (LEA, late-embryogenesis abundant protein) 基因有较高的同源性, 因此将其命名为 *AmLEA5*。

2.2 AmLEA5 蛋白二级结构和亚细胞定位预测

用 SOPMA 软件预测沙冬青 *AmLEA5* 基因所编码的蛋白二级结构。如图 2A、图 2B 所示, 该蛋白中 56 个氨基酸残基形成 α -螺旋结构, 含量为 57.14%, 主要集中在 N 端; 31 个氨基酸形成无规卷曲结构, 主要分布在 C 端, 占总数 31.63%。此外, 该蛋白中还有 3.06% 的伸展链和 8.16% 的 β -转角结构。使用 Predictprotein 软件通过多种方法验证了上述结果的可靠性。

```

ctctgattgactacgcaaattcatattaaagaataagcttccgctgtcataaaacca
ttcactataactcttctttcacacagtttcaattcaattcaattgaatcacta
120 atggctgcctccttcaactaacgccaaggttctctctgtctcgtt
    M A R S F T N A K V L S V L V
165 cttaatagattctcaagctcctcaccagacgtgggatgtatgcga
    L N R F S S S L T R R G Y A A
210 gcaagtgcacacaaagcgcaaaaagggaggagtgtctccatt
    A S A T Q S A K R G G V V S I
255 agcggcaagatggcacccaagtcaggggaagagaagaggccaca
    S G K M A P K S G E E K R A T
300 accgacaagtgctgtgggtgccagaccctgtgactgttactac
    T D K V S W V P D P V T G Y Y
345 aaacctgagaacattaagaaattgatgtctgaattgcgcgct
    K P E N I K E I D V A E L R A
390 acactcttgggaaaaatcaattaa 416
    T L L G K K F N *
    ttccaactgtgggtctcttcttcatgatcggagtcagatgact
    tgaacaaatctagagtattctgatgaaagacatggaatggaatattg
atgttgagatgcatgtaccaagaactgattcacttttatagttt
aagtgttattttcttcaaaaaaaaaaagttgtattttta
ttgtaccagatatttttttttttttcaatcaatag
aatgaccattattaagggtcaaaaaaaaaaaaaaaaaaaaaaaaaa

```

图 1 *AmLEA5* 的核苷酸序列及编码的氨基酸序列
Fig. 1 Nucleotide sequences and coded amino acid sequences of *AmLEA5*

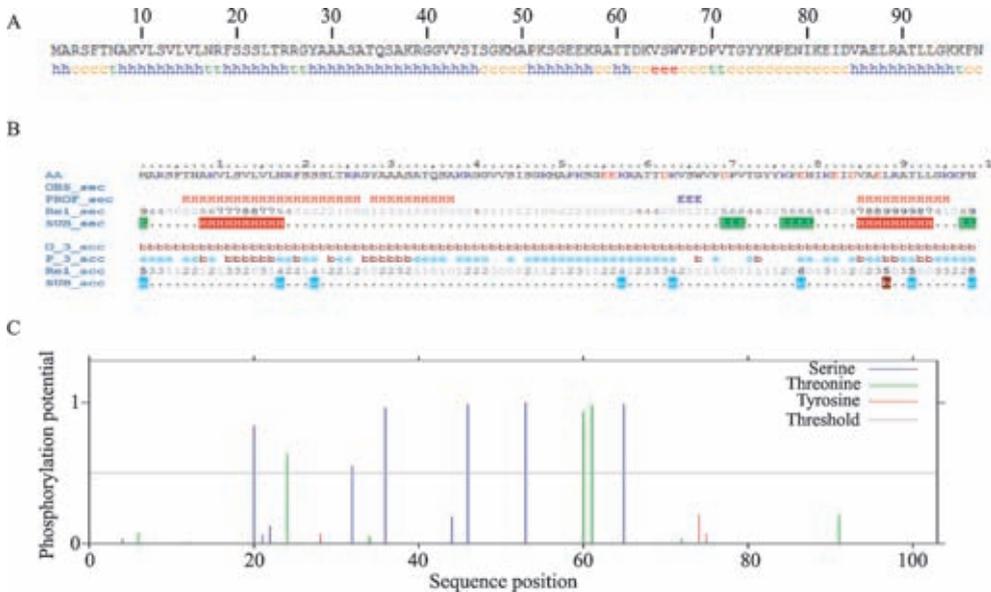
ProtParam 软件预测显示 *AmLEA5* 总平均疏水指数 (GRAVY) 为 -0.345; Predictprotein 软件多方法分析显示暴露氨基酸残基为 73.47% (图 2B), 以上结果都说明 *AmLEA5* 为亲水性蛋白。

用服务器 NetPhos 2.0 和 PROSITE 预测 *AmLEA5* 氨基酸序列的修饰位点发现该段序列中含有 1 个 cAMP-和 cGMP-依赖的蛋白激酶 (PKA 和 PKG) 磷酸化位点, 1 个酪蛋白激酶 II 磷酸化位点, 4 个蛋白激酶 C (PKC) 磷酸化位点 (图 2C)

经 Signal 3.0 预测发现 *AmLEA5* 的 N-端有线粒体定位信号, 切割位点是第 22 与第 23 氨基酸之间, 说明该蛋白可能主要位于线粒体。用 WoLF PSORT、SubLoc v1.0 和 Target P 等服务器预测发现该蛋白在沙冬青细胞中广泛存在, 可能在线粒体、叶绿体及其类囊体膜、细胞质中都有分布。

2.3 *AmLEA5* 的系统发生关系

用 BLAST 分析并用 MEGA 5.05 软件构建系统进化树 (图 3) 发现, *AmLEA5* 与豆科模式植物蒺藜苜蓿 *LEA* (ACJ84182.1) 的同源性高达 77%, 亲缘关系最近; 与其他植物中的 *LEA*, 如大豆 *LEA* (NP001237596.1)、拟南芥 *LEA3-like* (NP171781.1)、陆地棉 *LEA5D* (P46522.1)、烟草 *LEA5* (AAC06242.1)、柽柳 *LEA* (ABG54481.1)、西伯利亚白杨 *LEA5*

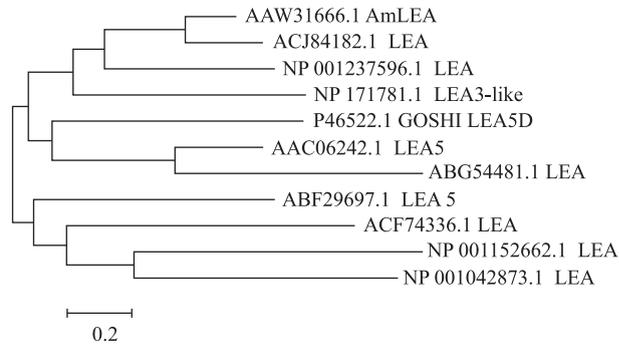


A; SOPMA 预测蛋白二级结构, h: α -螺旋, c: 无规卷曲, t: β -折叠, e: 延伸链; B; PredictProtein 预测蛋白二级结构和亲水性; sec: 二级结构; acc: 亲水性; H: α -螺旋; E: 延伸链; b: 埋藏的氨基酸残基; e: 暴露的氨基酸残基; C: NetPhos 2.0 服务器预测 *AmLEA5* 氨基酸序列磷酸化修饰位点
A; SOPMA; Secondary structure prediction, h: α -helix, c: Random coil, t: β -turn, e: Extended strand, B; PredictProtein server; Secondary structure and solvent accessibility prediction, sec: Secondary structure, acc: Solvent accessibility, H: α -Helix, E: Extended strand, b: Buried, e: Exposed, C; NetPhos 2.0 server; phosphorylation sites in sequence of *AmLEA5*

图 2 沙冬青 *AmLEA5* 二级结构、亲水性及磷酸化位点预测

Fig. 2 Secondary structure, hydrophobicity and phosphorylation sites of the coded protein *AmLEA5* of *A. mongolicus*

(ABF29697.1)、花生 LEA(ACF74336.1)、玉米 LEA(NP 001152662.1)、粳稻 LEA(NP 001042873.1)等相比同源性也较高。用 DNASTar 分析^[28]发现它们 N-末端和 C-末端的 30 个氨基酸都比较保守,而且大部分都属于 LEA5 蛋白,只有拟南芥 LEA3-like(NP 171781.1)属于 LEA3 蛋白。有报道称,LEA3 蛋白包含有部分 LEA5 蛋白^[29],基于以上分析和系统进化关系确定该沙冬青 LEA 为第 5 族 LEA 蛋白。



AAW31666.1:沙冬青 *Ammopiptanthus mongolicus*, ACJ84182.1:蒺藜
苜蓿 *Medicago truncatula*, NP 001237596.1:大豆 *Glycine max*,
NP 171781.1:拟南芥 *Arabidopsis thaliana*, P46522.1:陆地棉
Gossypium hirsutum, AAC06242.1:烟草 *Nicotiana tabacum*,
ABG54481.1:怪柳 *Tamarix androssowii*, ABF29697.1:西伯利亚
白杨 *Populus suaveolens*, ACF74336.1:花生 *Arachis hypogaea*,
NP 001152662.1:玉米 *Zea mays*, NP 001042873.1:粳稻
Oryza sativa Japonica Group

图 3 沙冬青 AmLEA5 与其他同源蛋白的系统进化分析

Fig. 3 Phylogenetic analysis of AmLEA5 in *A. mongolicus* and other homologous proteins

2.4 AmLEA5 密码子偏好性分析

ENc 值(effective number of codons)为基因的密码子偏好性程度提供了一个客观的评判标准,代表着特定基因中同义密码子非均衡使用的偏好程度。

该值的范围在 20(每个氨基酸只使用一个密码子的极端情况)到 61(各个密码子均被平均使用)之间,越靠近 20 偏好性越强。本研究应用 CHIPS 和 CodonW 在线程序计算沙冬青 AmLEA5 的 ENc 值、GC 含量和 GC3s 值分别为 57.785、48.82% 和 48.48%。该基因 ENc 值偏大,表明 AmLEA5 各密码子在编码氨基酸时出现的频率比较一致;AmLEA5 编码区 GC 含量适中,GC3s 比 GC 含量略低,表明沙冬青使用以 A、T 结尾的密码子略多,在整个编码区序列中 A + T 含量略大于 G + C。

相对同义密码子使用度(RSCU)是指对于某一特定的密码子在编码对应氨基酸的同义密码子中的相对概率。RSCU 值与氨基酸的使用及密码子的丰度无关,它能直观地反映出密码子使用的偏好性程度。如果密码子的使用没有偏好,则该密码子的 RSCU 值等于 1。当某一密码子的 RSCU 值大于 1,则表明该密码子的使用频率相对较高,反之亦然。Fraction 表示各个密码子在编码该氨基酸的密码子中所占的比例(各个比例相加总和 = 1)。Frequency 代表该密码子在编码基因总密码子中出现的频率,即在 1000 个密码子中出现的次数。CUSP 和 CodonW 在线程序计算结果(表 1)表明,在沙冬青 AmLEA5 的密码子中,25 个密码子的 RSCU 值大于 1,其中有 13 个密码子是以 A 或 T 碱基结尾的,有 12 个密码子是以 C 或 G 碱基结尾的,前者比后者多 1 个密码子。同样,Fraction 值和 Frequency 值较大的是 AmLEA5 基因的偏好密码子,从这两项指标中也可以看出该基因使用 A/T 结尾的密码子比使用 C/G 碱基结尾密码子略多。这就表明该基因密码子偏好性较平均,但是略偏好于使用 A/T 结尾的密码子。

表 1 CUSP 和 CodonW 在线程序分析 AmLEA5 基因密码子偏好性

Table 1 Condon bias of AmLEA5 gene analyzed by CUSP and CodonW program

密码子 Codon	氨基酸 AA	比例 Fraction	频率(%) Frequency	个数 Number	相对同义密码子使用度 RSCU
GCA	A	0.455	50.505	5	1.82
GCC	A	0.182	20.202	2	0.73
GCT	A	0.364	40.404	4	1.45
GAC	D	0.667	20.202	2	1.33
GAT	D	0.333	10.101	1	0.67
GAA	E	0.600	30.303	3	1.20
GAG	E	0.400	20.202	2	0.80
TTC	F	1.000	30.303	3	2.00
GGA	G	0.286	20.202	2	1.14

表 1(续)

密码子 Codon	氨基酸 AA	比例 Fraction	频率(%) Frequency	个数 Number	相对同义密码子使用度 RSCU
GGC	G	0.143	10.101	1	0.57
GCG	G	0.429	30.303	3	<u>1.71</u>
GGT	G	0.143	10.101	1	0.57
ATT	I	1.000	30.303	3	<u>3.00</u>
AAA	K	0.400	40.404	4	0.80
AAG	K	0.600	60.606	6	<u>1.20</u>
CTC	L	0.429	30.303	3	<u>2.57</u>
CTT	L	0.429	30.303	3	<u>2.57</u>
TTC	L	0.143	10.101	1	0.86
ATG	M	1.000	20.202	2	1.00
AAC	N	0.500	20.202	2	1.00
AAT	N	0.500	20.202	2	1.00
CCA	P	0.250	10.101	1	1.00
CCC	P	0.250	10.101	1	1.00
CCT	P	0.500	20.202	2	<u>2.00</u>
CAA	Q	1.000	10.101	1	<u>2.00</u>
AGA	R	0.286	20.202	2	<u>1.71</u>
AGG	R	0.286	20.202	2	<u>1.71</u>
CGC	R	0.286	20.202	2	<u>1.71</u>
CGT	R	0.143	10.101	1	0.86
AGC	S	0.273	30.303	3	<u>1.64</u>
AGT	S	0.091	10.101	1	0.55
TCA	S	0.182	20.202	2	<u>1.09</u>
TCC	S	0.273	30.303	3	<u>1.64</u>
TCG	S	0.091	10.101	1	0.55
TCT	S	0.091	10.101	1	0.55
ACA	T	0.429	30.303	3	<u>1.71</u>
ACC	T	0.286	20.202	2	<u>1.14</u>
ACT	T	0.286	20.202	2	<u>1.14</u>
GTC	V	0.111	10.101	1	0.44
GTG	V	0.333	30.303	3	<u>1.33</u>
GTT	V	0.556	50.505	5	<u>2.22</u>
TGG	W	1.000	10.101	1	1.00
TAC	Y	0.667	20.202	2	<u>1.33</u>
TAT	Y	0.333	10.101	1	0.67
TAA	*	1.000	10.101	1	0

A: 丙氨酸, D: 天冬氨酸, E: 谷氨酸, F: 苯丙氨酸, G: 甘氨酸, I: 异亮氨酸, K: 赖氨酸, L: 亮氨酸, M: 甲硫氨酸, N: 天冬酰胺, P: 脯氨酸, Q: 谷氨酰胺, R: 精氨酸, S: 丝氨酸, T: 苏氨酸, V: 缬氨酸, W: 色氨酸, Y: 酪氨酸。下划线表示大于 1 的 RSCU 值。*: 无氨基酸

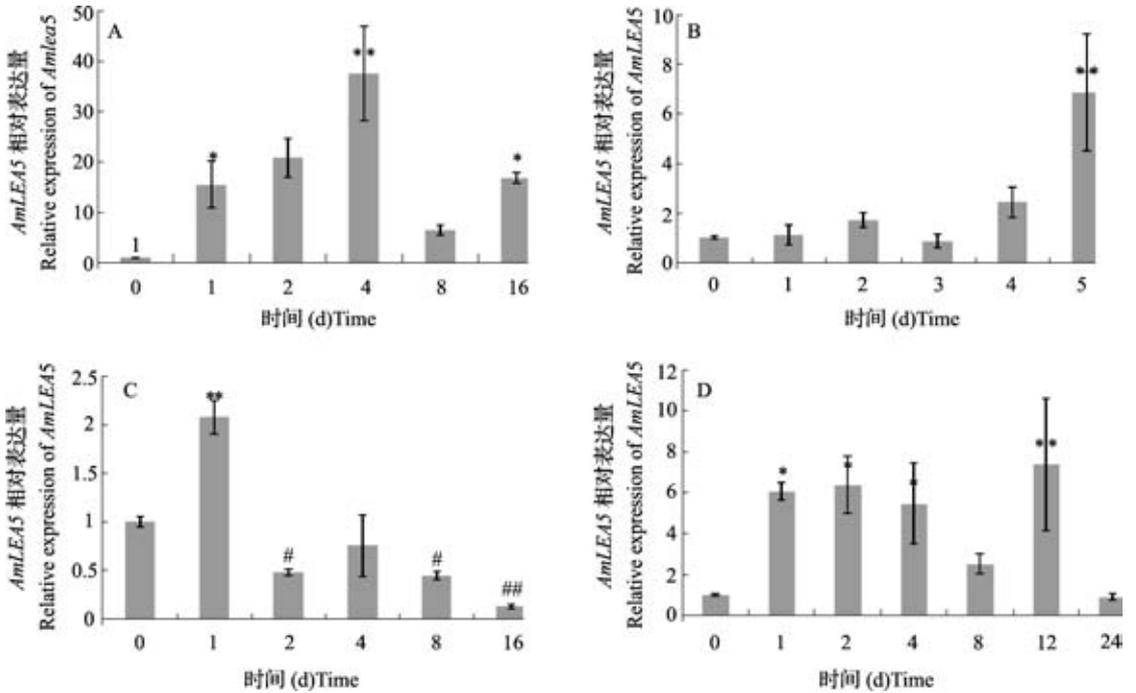
A: Ala, D: Asp, E: Glu, F: Phe, G: Gly, I: Ile, K: Lys, L: Leu, M: Met, N: Asn, P: Pro, Q: Gln, R: Arg, S: Ser, T: Thr, V: Val, W: Trp, Y: Tyr.

The data with underline mean that the value of RSCU > 1. *: No amino acid

2.5 *AmLEA5* 对非生物胁迫的应答反应

实时荧光定量 PCR 的结果显示:*AmLEA5* 基因在低温、干旱、盐和热胁迫条件下的表达特性是不同的(图4)。在这4种胁迫条件下,*AmLEA5* 的相对表达量都出现了明显的升高。在低温条件下(图4A)随着处理时间的延长,基因的相对表达量逐渐升高,第4天达到最高,随后骤然下降,但是低温处理的第1天到第16天内基因的相对表达量始终高于对照

组。干旱条件下(图4B),沙冬青在处理的前4 d *AmLEA5* 基因相对表达量没有明显升高,到第5天相对表达量明显升高,是对照组的6倍多。盐处理条件下(图4C)基因相对表达量在第1天达到最高,随后的第2天到第16天均低于对照组。热处理条件下(图4D),1~12 h内 *AmLEA5* 表达量均明显高于对照组,只是在第8小时,相对表达量明显下降;到第24小时相对表达量低于对照组。



A:4 °C 处理;B:干旱处理;C:NaCl 处理;D:40 °C 处理

*: $P < 0.05$; **: $P < 0.01$;代表胁迫后相对表达量高于胁迫前(0)相对表达量

#: $P < 0.05$;##: $P < 0.01$;代表胁迫后相对表达量低于胁迫前(0)相对表达量

A:4 °C treatment, B:Drought treatment, C:NaCl treatment, D:40 °C treatment, *: $P < 0.05$, **: $P < 0.01$. Stand for the relative quantity of stresses higher than of pretreatment (0), #: $P < 0.05$, ##: $P < 0.01$. Stand for the relative quantity of stresses lower than of pretreatment(0)

图4 *AmLEA5* 在低温、干旱、盐和热胁迫下的表达模式

Fig. 4 Expression profiles of *AmLEA5* under cold, drought, salt and heat stresses

2.6 YFP-*AmLEA5* 融合蛋白的亚细胞定位

Gateway™ 重组技术是一种高效的大规模克隆系统,而且对载体和宿主没有依赖性^[30]。本试验利用这种快速而可靠的克隆技术构建双元表达载体,用基因枪将双元表达载体转化到洋葱内表皮,在激光共聚焦显微镜下观察 YFP-*AmLEA5* 融合蛋白在洋葱内表皮细胞中的定位,这种定位方法稳定,易于观察。如图5, YFP-*AmLEA5* 主要定位在细胞核、细胞质及细胞膜。

3 讨论

LEA 蛋白是一类重要的抗逆蛋白,目前国内外

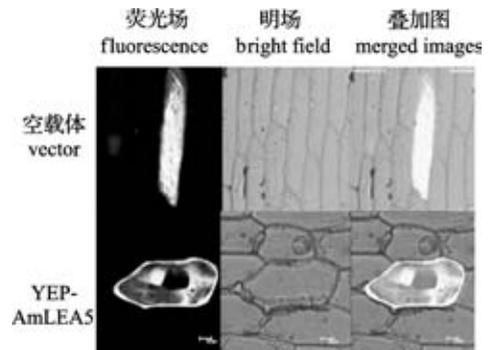


图5 YFP-*AmLEA5* 融合蛋白在洋葱表皮细胞中的瞬时表达模式

Fig. 5 The transient expression patterns of YFP-*AmLEA5* fusion protein in onion epidermis cells

关于 *LEA* 基因与非生物胁迫之间的研究日益增多。研究发现 *LEA* 蛋白具有稳定细胞膜、清除活性氧自由基、结合金属离子等功能^[31],事实上,*LEA* 的功能主要是通过保护过氧化氢酶、乳酸脱氢酶等的活性在细胞内发挥防御作用的,而且发现 *LEA* 的保护作用明显高于脯氨酸和甜菜碱^[32]。

LEA 蛋白的结构决定它的性质和功能,生物信息学分析表明,沙冬青 *AmLEA5* 属于第 5 族 *LEA*,热稳定性不高,但亲水性很强。这是由于基元序列富含 α -螺旋结构,在胁迫条件下可能形成高度螺旋折叠,该结构具有稳定脂膜或功能蛋白的作用,可以阻止水分大量散失,从而将外界环境对细胞内代谢的影响降到最低。因而,*AmLEA5* 对于非生物胁迫导致的细胞失水具有防御和抵抗作用。根据已有研究得知,信号传导在细胞代谢和植物生长过程中起到至关重要的作用,蛋白磷酸化和去磷酸化参与细胞信号传导^[33-34],本研究对 *AmLEA5* 氨基酸序列修饰位点进行预测发现该序列中含有多个磷酸化位点,但是这些磷酸化位点的磷酸化与去磷酸化是否会对低温等胁迫信号的传导起作用,以及如何调节细胞信号传导目前尚不清楚,还有待进一步研究。

LEA 在细胞中的分布对植物抗逆性也有重要影响。大部分 *LEA* 蛋白没有明显的跨膜结构,在细胞内的分布极为广泛,可以存在于细胞核、线粒体^[35]、叶绿体^[36]或细胞质中。软件预测结果显示,*AmLEA5* 可能存在于细胞质、线粒体或叶绿体中。而 YFP-*AmLEA5* 融合蛋白的亚细胞定位表明,*AmLEA5* 主要分布于细胞核、细胞质及细胞质膜上,线粒体和叶绿体中是否存在 *AmLEA5* 仍无法确定。由于 *AmLEA5* 富含 α -螺旋结构,所以位于质膜上的 *AmLEA5* 可以与膜结合,稳定细胞膜,从而阻止水分进一步散失,保证细胞内各物质生长代谢所需的适宜环境。

AmLEA5 的系统发生分析结果显示,沙冬青 *AmLEA5* 和蒺藜苜蓿 (*Medicago truncatula* L.) ACJ84182.1 的亲缘关系最近,它们同属于豆科植物,根系发达,共性较多。此外,系统进化关系也可以通过密码子偏好性来分析^[23]。

密码子偏好性的研究便于人们了解物种与密码子使用之间的规律和特点。研究表明,单子叶植物和双子叶植物的密码子偏好性存在明显差异^[37],双子叶植物 GC 含量低于单子叶植物,即双子叶植物更倾向于使用 A 或 T 结尾的密码子。通常,G 或 C 结尾的密码子更容易导致基因突变,因此,单子叶植

物如禾本科植物比双子叶植物更容易发生变异^[38]。密码子偏好性分析发现 *AmLEA5* 基因密码子偏好性较平均,但是略偏好于使用以 A 或 T 结尾的密码子,基因稳定性较好。沙冬青属于豆科双子叶植物,*AmLEA5* 基因的密码子偏好性分析结果符合双子叶植物 GC 含量低的事实。此外,密码子偏好性的研究,也为转基因技术提供理论支持,如果采用转基因手段深入了解 *AmLEA5* 的功能,将该基因转入其他双子叶植物中更有助于提高试验的成功率和可靠性,作为模式植物的烟草、拟南芥都可以作为候选受体植物。

有研究表明,许多 *LEA* 蛋白可以提高植物抗冻能力。小麦低温诱导脱水素蛋白(WCOR410)在遭受低温胁迫时可以在植物组织中大量积累^[39];低温诱导拟南芥中的冷调节脱水素(COR47)^[12]大量表达,可以提高植株抗寒性。此外,X. Zhao 等^[40]研究发现,紫杆怪柳 *LEA* 基因的过表达可以增强转基因蓝莓对 0 °C 低温的抗性。本研究表明,*AmLEA5* 基因的相对表达量受低温诱导后,升高最明显,持续时间最长,由此推知,*AmLEA5* 基因除了在抗旱、耐盐、耐热方面发挥作用外,可能主要参与低温伤害的防御机制。将低温诱导 *AmLEA5* 基因相对表达量升高的结果与生物信息学结果结合起来分析,推测正是由于自由 α -螺旋结构的存在使得 *AmLEA5* 蛋白可以与胞内对低温等敏感的酶或蛋白相互作用,使其结构稳定;或者可能给细胞内的束缚水提供一种结合界面,脱水时细胞结构不致遭受更大的破坏。

抵抗非生物胁迫是多种蛋白共同参与,协同作用的结果,因此深入开展沙冬青 *AmLEA5* 基因及相关基因功能和机理的研究,不仅可以揭示沙冬青抗逆性分子机理,保护濒危珍稀物种,而且可以为抗逆育种提供优良基因。

参考文献

- [1] Dure L, Greenway S C, Galau G A, et al. Developmental biochemistry of cottonseed embryogenesis and germination: changing messenger ribonucleic acid populations as shown by in vitro and in vivo protein synthesis[J]. *Biochemistry*, 1981, 20(14): 4162-4168
- [2] Ried J L, Walker-Simmons M K. Group 3 late embryogenesis abundant proteins in desiccation-tolerant seedlings of wheat (*Triticum aestivum* L.) [J]. *Plant Physiol*, 1993, 102(1): 125-131
- [3] Hsing Y C, Chen Z Y, Shih M D, et al. Unusual sequences of group 3 *LEA* mRNA inducible by maturation or drying in soybean seeds[J]. *Plant Mol Biol*, 1995, 29(4): 863-868
- [4] Campbell S A, Crone D E, Ceccardi T L, et al. A ca. 40 kDa maize (*Zea mays* L.) embryo dehydrin is encoded by the *dhn2* locus on chromosome 9[J]. *Plant Mol Biol*, 1998, 38(3): 417-423
- [5] Browne J, Tunnacliffe A, Burnell A, et al. Anhydrobiosis: Plant

- desiccation gene found in a nematode[J]. *Nature*,2002,416:38
- [6] 何军贤,傅家瑞. 种子 *Lea* 蛋白的研究进展[J]. *植物生理学通讯*,1996,32(4):241-246
- [7] Puhakainen T, Hess M W, Makela P, et al. Overexpression of multiple dehydrin genes enhances tolerance to freezing stress in *Arabidopsis*[J]. *Plant Mol Biol*,2004,54(5):743-753
- [8] Kim H S, Lee J H, Kim J J, et al. Molecular and functional characterization of *CaLEA6*, the gene for a hydrophobic LEA protein from *Capsicum annum*[J]. *Gene*,2005,344(3):115-123
- [9] 杨天旭,汪耀富,宋世旭,等. 逆境胁迫下植物 LEA 蛋白的研究进展[J]. *干旱地区农业研究*,2006,24(6):120-124
- [10] Dure L. A repeating 11-mer amino acid motif and plant desiccation[J]. *Plant J*,1993,3(3):363-369
- [11] Tolleter D, Jaquinod M, Mangavel C, et al. Structure and function of a mitochondrial late embryogenesis abundant protein are revealed by desiccation[J]. *Plant Cell*,2007,19(5):1580-1589
- [12] Hundertmark M, Hinch D K. LEA (late embryogenesis abundant) proteins and their encoding genes in *Arabidopsis thaliana*[J]. *BMC Genomics*,2008,9:118
- [13] Dorit N, Gozal B H, Yuval E, et al. Drought, heat and salt stress induce the expression of a citrus homologue of an atypical late-embryogenesis *Lea5* gene[J]. *Plant Mol Biol*,1995,27:619-622
- [14] 林清芳,王茅雁,刘佳杰,等. 沙冬青细胞与分子生物学研究进展[J]. *植物遗传资源学报*,2010,11(6):793-797
- [15] 陶玲,李新荣,刘新民,等. 中国珍稀濒危荒漠植物保护等级的定量研究[J]. *林业科学*,2001,37(1):52-57
- [16] 田晓明,陈世忠,屠鹏飞,等. 沙冬青地上部分的化学成分研究[J]. *中国中药杂志*,2008,33(19):2204-2206
- [17] Liu M Q, Shen X, Lu C F, et al. Functional analysis of cold inducible cDNA clones in the legume *Ammiptanthus mongolicus*[J]. *Cryo Letters*,2005,26(4):213-222
- [18] Liu M Q, Lu C F, Yin W L, et al. Characterization and functional analysis of a cold-induced *AmCIP* gene encoding a dehydrin-like protein in *Ammiptanthus mongolicus*[J]. *DNA Sequence*,2006,17(5):342-349
- [19] 刘瑞玲,刘美芹,史军娜,等. 过量表达沙冬青疏基蛋白酶抑制剂基因 *AmPI* 提高大肠杆菌低温与热胁迫抗性[J]. *应用与环境生物学报*,2010,16(3):341-346
- [20] Liu R L, Liu M Q, Liu J, et al. Heterologous expression of a *Ammiptanthus mongolicus* late embryogenesis abundant protein gene (*AmLEA*) enhances *Escherichia coli* viability under cold and heat stress[J]. *Plant Growth Regul*,2010,60:163-168
- [21] 刘美芹,尹伟伦,卢存福,等. 沙冬青抗寒性分子基础研究[M]. 北京:中国环境科学出版社,2009
- [22] 李平,白云凤,冯瑞云,等. 籽粒苋苹果酸酶(*NAD-ME*)基因密码子偏好性分析[J]. *应用与环境生物学报*,2011,17(1):12-17
- [23] Li M L, Zhao Z Y, Cai M S, et al. Characterization of synonymous codon usage bias in the pseudorabies virus *US₁* gene[J]. *Virologica Sinica*,2012,27(5):303-315
- [24] 王艳萍,刘美芹,卢存福,等. 沙冬青热胁迫相关蛋白基因 *AmHsa32* 超表达提高大肠杆菌的抗热性[J]. *北京林业大学学报*,2012,34(5):37-43
- [25] Shi J, Liu M Q, Lu C F, et al. Reference gene selection for qPCR in *Ammiptanthus mongolicus* under abiotic stresses and expression analysis of seven ROS-scavenging enzyme genes[J]. *Plant Cell Rep*,2012,31(7):1245-1254
- [26] 师静,刘美芹,卢存福,等. 沙冬青胚胎晚期发生丰富蛋白基因序列及表达特性分析[J]. *北京林业大学学报*,2012,34(4):114-119
- [27] 张岳平,瞿华香. 新模式植物短柄二叶草 *SGT1* 和 *RARI* 基因沉默和蛋白纯化载体构建[J]. *植物遗传资源学报*,2011,12(1):138-144
- [28] Thompson J D, Higgins D G, Gibson T J, et al. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice[J]. *Nucleic Acids Res*,1994,22:4676-4680
- [29] Wise M J. LEAping to conclusions: A computational reanalysis of late embryogenesis abundant proteins and their possible roles[J]. *BMC Bioinformatics*,2003,4:52-70
- [30] 梅文倩,宋文强,朱玉贤,等. 利用 Gateway 克隆技术大规模克隆拟南芥转录因子[J]. *分子植物育种*,2004,2(3):358-364
- [31] Kosová K, Vitámvás P, Prášil I T, et al. Expression of dehydrins in wheat and barley under temperatures[J]. *Plant Sci*,2011,180(1):46-52
- [32] Hara M, Terashima S, Kuboi T, et al. Characterization and cryoprotective activity of cold-responsive dehydrin from *Citrus unshiu*[J]. *J Plant Physiol*,2001,158(10):1333-1339
- [33] 孔令安,汪矛,林金星,等. 蛋白质可逆磷酸化调节植物细胞离子跨膜运动研究进展[J]. *西北植物学报*,2008,28(7):1491-1499
- [34] 马雅琴,翁跃进,赵勇,等. 植物耐盐相关基因克隆的研究进展[J]. *植物遗传资源学报*,2004,5(1):81-86
- [35] Borovskii G B, Stupnikova I V, Antipina A I, et al. Accumulation of dehydrin-like proteins in the mitochondria of cereals in response to cold, freezing, drought and ABA treatment[J]. *BMC Plant Biol*,2002,2:5-10
- [36] Mueller J K, Heckathorn S A, Fernando D, et al. Identification of a chloroplast dehydrin in leaves of mature plants[J]. *Plant Sci*,2003,164(4):535-542
- [37] Murray E E, Lotzer J, Eberle M, et al. Codon usage in plant genes[J]. *Nucl Acids Res*,1989,17(2):477-498
- [38] Kawabe A, Miyashita N T. Patterns of codon usage bias in three dicot and four monocot plant species[J]. *Genes Genet Syst*,2003,78(5):343-352
- [39] Houde M, Hallaire S, Dong N D, et al. Overexpression of the acidic dehydrin WCOR410 improves freezing tolerance in transgenic strawberry leaves[J]. *Plant Biotechnol J*,2004,2(5):381-387
- [40] Zhao X, Zhan L P, Zou X Z, et al. Improvement of cold tolerance of the half-high bush Northland blueberry by transformation with the *LEA* gene from *Tamarix androssowii*[J]. *Plant Growth Regul*,2011,63:13-22