

刘 丹,毕德浣,陈 鑫,等. 狗枣猕猴桃胚胎发育晚期富集蛋白的生物信息学分析[J]. 江苏农业科学,2019,47(24):30-33.
doi:10.15889/j.issn.1002-1302.2019.24.007

狗枣猕猴桃胚胎发育晚期富集蛋白的生物信息学分析

刘 丹,毕德浣,陈 鑫,王立凤

(牡丹江师范学院,黑龙江牡丹江 157012)

摘要:胚胎发育晚期富集蛋白(LEA)是目前最受关注的逆境胁迫响应蛋白,在植物受到多种胁迫时 LEA 蛋白大量表达,从而减轻了植物受到的逆境伤害。利用生物信息学方法对狗枣猕猴桃胚胎发育晚期富集蛋白(AkLEA)进行结构及功能预测。结果表明:狗枣猕猴桃 AkLEA 的 cDNA 全长为 1 292 bp,CDS 为 657 bp,编码蛋白由 221 个氨基酸构成、具有 AkLEA 的保守序列,是 LEA 蛋白(PLN03160)超家族的一员,是一个亲水性、稳定的蛋白。与其他 17 种植物的 LEA 氨基酸序列进行对比,结果表明,狗枣猕猴桃 AkLEA 基因与中华猕猴桃变种亲缘关系较近,而与杨树、石榴等植物亲缘关系较远。本研究结果可为狗枣猕猴桃 AkLEA 基因功能的进一步研究提供参考。

关键词:狗枣猕猴桃;AkLEA 蛋白;生物信息学

中图分类号:S663.401 **文献标志码:**A **文章编号:**1002-1302(2019)24-0030-04

狗枣猕猴桃(*Actinidia kolomikta*)别名狗枣子,是猕猴桃科(Actinidiaceae)猕猴桃属(*Actinidia*)多年生落叶藤本植物,是最为耐寒的猕猴桃种类之一。主要分布于我国东三省、四川、云南等高山地区,俄罗斯、日本及朝鲜半岛也有分布。喜生于阔叶、针叶混交灌木林且通风良好、土质湿润、腐殖质肥沃的自然环境中。其果实富含维生素 C、有机酸、蛋白质、多种糖类及人体所需的多种微量元素和氨基酸。狗枣猕猴桃的抗逆性强,还具有药物开发、食用保健等经济效益。是东北地区蕴藏丰富的野生浆果资源,是珍贵的抗寒种质资源^[1-2]。胚胎发育晚期富集蛋白(late embryogenesis abundant protein, LEA),分子量在 10~30 ku 之间,具有高亲水性,属于逆境胁迫响应蛋白^[3-4]。1989 年 Dure 等在棉花种子中首次发现 LEA 的存在,后在大豆、水稻、拟南芥、葡萄等植物中陆续发现^[5],LEA 的结构和功能一直以来都吸引着众多科学家的兴趣和重视,广泛存在于植物中。到目前为止已克隆获得上百种 LEA 基因^[6],通过各种分子生物学手段,人们发现,植物营养组织遭受干旱、低温、ABA 处理及盐胁迫等逆境胁迫后,会产生许多新的 mRNA 和蛋白质,该蛋白质中相当大一部分均与 LEA 蛋白有同源性^[7]。众多研究表明 LEA 基因的功能主要与抗逆性有关,植物营养组织遭受逆境胁迫后均能诱导 LEA 基因的表达^[8]。许多 LEA 基因转化试验也表明,LEA 可显著提高植物的抗逆性。LEA 蛋白富含赖氨酸、甘氨酸、组氨酸等极性氨基酸,缺乏色氨酸和半胱氨酸,无明显的二级结构,以一种无序状态存在于溶液中。目前,LEA 基因的研究越来越广泛,大多研究集中在植物抗逆性方面^[9]。鉴于 LEA 蛋

白在植物抗逆性方面的重要作用,本研究以狗枣猕猴桃转录组数据为参考,筛选到 LEA 蛋白的 cDNA 序列,利用生物信息学方法对枣猕猴桃胚胎发育晚期富集蛋白的理化性质、疏水性/亲水性、蛋白结构功能域、二级结构、三级结构及同源性等方面进行预测和分析,以期对狗枣猕猴桃抗寒能力形成的相关研究提供参考。

1 材料与方法

1.1 材料来源

狗枣猕猴桃胚胎发育晚期富集蛋白 cDNA 序列从笔者所在实验室转录组数据中获得;山黄麻(*Parasponia andersonii*)、东方苔草(*Trema orientale*)、欧洲栓皮栎(*Quercus suber*)、荷花(*Nelumbo nucifera*)、中华猕猴桃原变种(*Actinidia chinensis* var. *chinensis*)、苹果(*Malus domestica*)、土瓶草(*Cephalotus follicularis*)、博落回(*Macleaya cordata*)、可可树(*Theobroma cacao*)、甜椒(*Capsicum annuum*)、巴卡廷辣椒(*Capsicum baccatum*)、番茄(*Solanum lycopersicum*)、向日葵(*Helianthus annuus*)、香青蒿(*Artemisia annua*)、杨树(*Populus trichocarpa*)、石榴(*Punica granatum*)等植物氨基酸序列在 NCBI 数据库中下载。

1.2 方法

狗枣猕猴桃 AkLEA 的理化性质用 Prot Param(<https://web.expasy.org/cgi-bin/protparam/protparam>)进行分析;NCBI-CDD 进行保守结构域及蛋白家族预测(<https://www.ncbi.nlm.nih.gov/Structure/cdd/>);Inter-Prot Scan(<http://www.ebi.ac.uk/interpro/search/sequence-search>)预测功能位点;NetNGlyc(<http://www.cbs.dtu.dk/services/NetNGlyc/>)对 N-糖基化位点进行预测;NetOGlyc(<http://www.cbs.dtu.dk/services/NetOGlyc/>)对 O-糖基化位点进行预测;采用 Predict Protein 数据库(<https://www.predictprotein.org/>)对 AkLEA 蛋白进行二级结构预测;使用 SWISS-MODEL(<http://swissmodel.expasy.org/>)在线软件预测蛋白质三级

收稿日期:2019-03-12

基金项目:牡丹江师范学院科研项目(编号:QN2019005);黑龙江省教育厅项目(编号:1354MSYYB007);黑龙江药用植物开发与利用研究(编号:1353PT007);黑龙江省大学生创新创业训练计划(编号:201910233007)。

作者简介:刘 丹(1982—),女,河南省洛阳人,硕士,实验师,主要从事植物病理学研究。E-mail:swxld1@126.com。

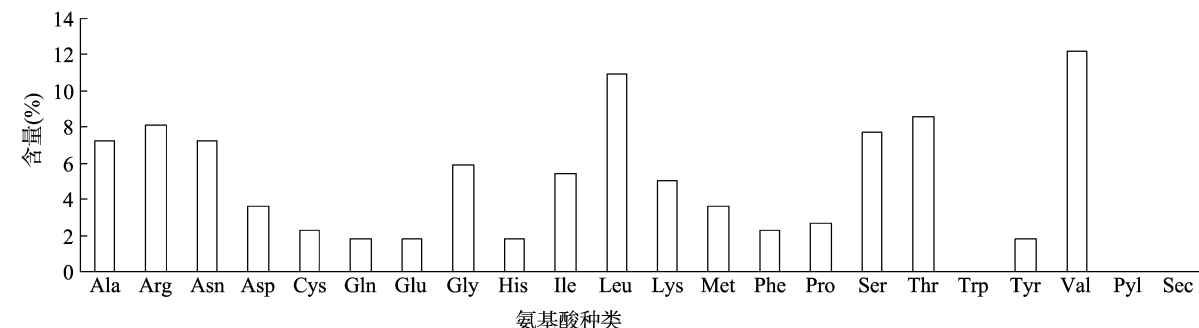
结构。从 NCBI (<https://www.ncbi.nlm.nih.gov/>) 数据库中
获得 19 条不同植物的 LEA 氨基酸序列信息,利用 MEGA 5.0
软件的 Neighbor-joining 法构建蛋白质系统进化树。

2 结果与分析

2.1 LEA 理化性质预测

利用 Prot Param 软件对狗枣猕猴桃 AkLEA 氨基酸序列
进行分析,结果表明,狗枣猕猴桃 AkLEA 基因 cDNA 全长为
1 292 bp,编码区(CDS)为 cDNA 上第 247~912 个碱基,共编
码 221 个氨基酸,分子量 24 130.33 u,理论等电点为 10.40,
负电荷残基总数(Asp + Glu)为 49,正电荷残基总数(Arg +

Lys)为 12,分子式为 $C_{1052}H_{1770}N_{314}O_{306}S_{13}$,总原子数 3 455,亲
水性(GRAVY)的平均水平:0.152。不稳定指数(instability
index, II)为 33.97,说明该蛋白较为稳定。该蛋白中含量最
多的为亮氨酸,占 12.2%,其次为异亮氨酸和苏氨酸,各占
10.9%和 8.6%。含量最少的为谷氨酰胺、谷氨酸、组氨酸、
酪氨酸,均为 1.8%(图 1)。有研究表明,大部分植物中的
LEA 蛋白富含赖氨酸、甘氨酸、组氨酸等极性氨基酸,缺乏色
氨酸和半胱氨酸,本研究狗枣猕猴桃色氨酸缺乏,但是存在
半胱氨酸,这是狗枣猕猴桃不同于已研究的其他植物的一个
特性。



Ala—丙氨酸; Arg—精氨酸; Asn—天冬酰胺; Asp—天冬氨酸; Cys—半胱氨酸; Gln—谷氨酰胺; Glu—谷氨酸; Gly—甘氨酸;
His—组氨酸; Ile—异亮氨酸; Leu—亮氨酸; Lys—赖氨酸; Met—蛋氨酸; Phe—苯丙氨酸; Pro—脯氨酸; Ser—丝氨酸; Thr—苏氨酸;
Trp—色氨酸; Tyr—酪氨酸; Val—缬氨酸; Pyl—吡咯赖氨酸; Sec—硒半胱氨酸

图1 狗枣猕猴桃 AkLEA 基因编码蛋白质的氨基酸组成

在自然界中常见的氨基酸均具有亲水性和疏水性,其排
列顺序侧链基因之间的相互作用决定了其合成蛋白质三维空
间结构的主要原因,通过测定和分析可以为高级结构的预测
提供参考。通过 Prot Scale 软件分析预测,结果表明:该基因

蛋白的整个肽链中均含有亲水性和疏水性氨基酸,得分最大
值为 3.867,在 56、57、58 个氨基酸处,最小值为 -2.600,在第
36 个氨基酸处。根据蛋白质亲水和疏水的得分判定该蛋白
为亲水蛋白(图 2)。

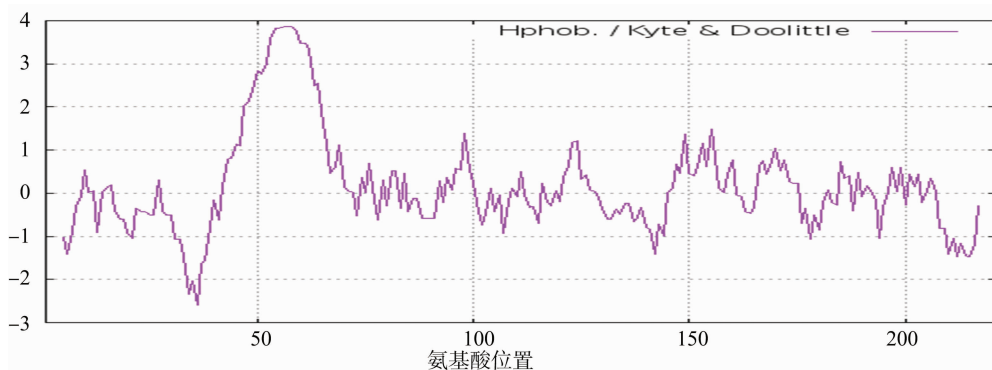


图2 AkLEA 蛋白亲水/疏水性分析

2.2 AkLEA 信号肽和跨膜结构

有些蛋白质合成以后,需要转移到特定位置发挥作用,信
号肽就是引导蛋白质到达特定位置的特殊氨基酸序列,一般
位于蛋白质的 N 端,长约 10~30 个氨基酸,包含信号肽的 C
端、N 端和疏水核心区。SignalP-4.1 预测 AkLEA 蛋白不含
信号肽,也不存在跨膜结构(图 3)。

2.3 狗枣猕猴桃 AkLEA 功能位点和结构域预测

通过氨基酸序列的比对,对狗枣猕猴桃 AkLEA 保守序列
及蛋白家族进行预测,结果表明,狗枣猕猴桃 AkLEA 是 LEA
基因家族的 LEA2 成员,它的第 100~200 氨基酸之间的序列
区域为 LEA2 组的保守结构域(图 4)。NCBI-CDD 和

Inter-ProtScan 预测结果表明,67~221 氨基酸为狗枣猕猴桃
AkLEA 的 N 端序列,1~39 氨基酸为狗枣猕猴桃 AkLEA 的 C
端序列。

2.4 AkLEA 二级、三级结构预测

用 Predictprotein 对 AkLEA 的二级结构进行预测,结果表
明,组成 AkLEA 的 221 个氨基酸中,其中 8.14% 的氨基酸可
能会形成螺旋(helix),43.44% 的氨基酸可能形成延伸链
(strand),48.42% 的氨基酸可能形成环(loop),无二硫键的形
成,没有特殊的二级结构。用 SWISS-MODEL 对 AkLEA 蛋
白进行三级结构的预测,结果如图 5 所示。

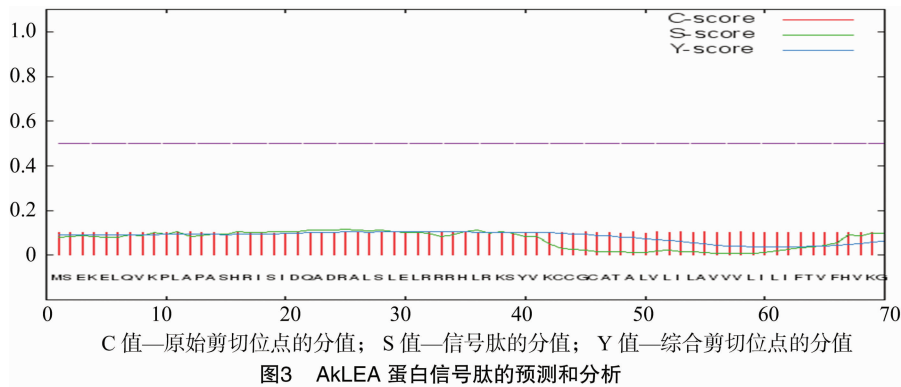


图3 AkLEA 蛋白信号肽的预测和分析

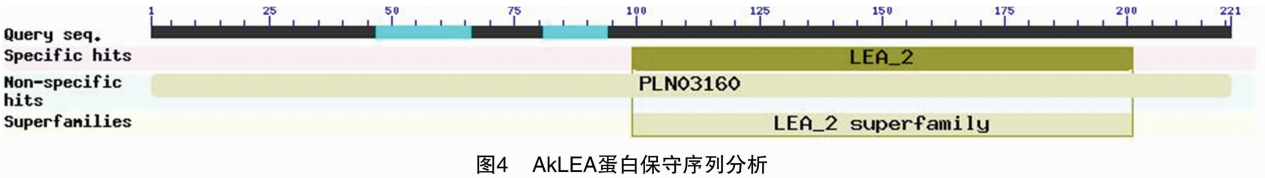


图4 AkLEA蛋白保守序列分析

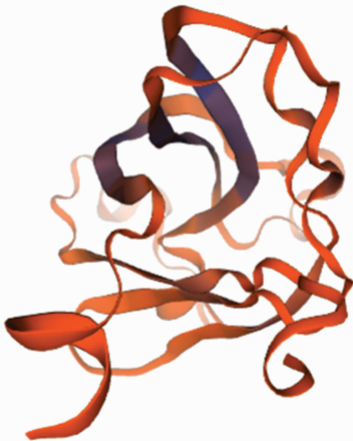


图5 AkLEA 蛋白的三级结构预测

2.5 蛋白质修饰位点的预测

蛋白质磷酸化是调节和控制蛋白质活力和功能的最基本、最普遍,也是最重要的机制。通过 predictprotein 分析表明,AkLEA 上含有 5 个糖基化位点(c N - glycosylation site)、7 个蛋白激酶 C 磷酸化位点(Protein kinase C phosphorylation site)、1 个酪蛋白磷酸激酶 II 磷酸化位点(Casein kinase II phosphorylation site)、5 个十四烷酰化位点(N - myristoylation site)(表 1)。

蛋白质的 N - 糖基化修饰是生物体调控蛋白质在组织和细胞中的定位、功能、活性、寿命和多样性的一种普遍的翻译后修饰。通过 NetNGlyc1.0 server 对 N - 糖基化位点进行预测,结果表明,该蛋白内有 5 个潜在的 N - 糖基化位点,分别在 93(NLTV)、146(NMTI)、172(NMSS)、198(NCTM)及 204(NATS)氨基酸处。O - 糖基化是另一种重要的蛋白质翻译

表 1 AkCHS1 蛋白修饰位点的预测

功能	位点
糖基化	93:NLTV;146:NMTT;172:NMSS;198:NCTM;204:NATS
蛋白激酶磷酸化	2;SEK;15;SHR;101;SVK;109;SFK;142;TLR;178;SVR;194;TVR
酪蛋白磷酸激酶 II 磷酸化	2;SEKE
十四烷酰化	45;GCATAL;82;GLDLSN;91;GGNLTV;124;GVAUGE;135;GVARAR

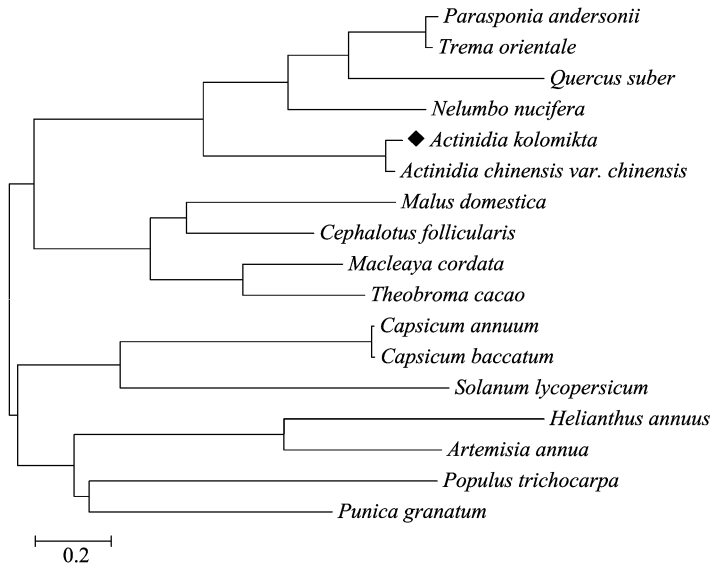
后修饰,O - 糖链能维持所连接蛋白质部分的空间构象,NetOGlyc4.0 server 对 O - 糖基化位点进行预测,预测分数均大于 0.5 的认为是阳性,故推测该蛋白内含 3 个 O - 糖基化位点,分别在 22、26、108 氨基酸处。

2.6 氨基酸序列同源性对比

在 NCBI 数据库中分别下载山黄麻、东方苔草、欧洲栓皮栎、荷花、中华猕猴桃原变种、苹果、土瓶草、博落回、可可树、甜椒、巴卡廷辣椒、番茄、向日葵、香青蒿、杨树、石榴等 17 种植物的 LEA 序列,与狗枣猕猴桃 AkLEA 氨基酸序列进行同源性分析,结果表明,狗枣猕猴桃 AkLEA 与中华猕猴桃原变种亲缘关系较近,而与杨树、石榴等植物亲缘关系较远(图 6)。

3 讨论

本研究利用已测的狗枣猕猴桃转录组数据,筛选出狗枣猕猴桃 AkLEA 基因 cDNA 序列信息,并对其进行了生物信息学分析。结果表明,AkLEA 基因 cDNA 总长 1 292 bp,CDS 为 657 bp,编码蛋白由 221 个氨基酸构成。具有 LEA 的保守序列,是 LEA2(PLN03160)超家族的一员,是一个亲水性、稳定的蛋白。LEA_2 组中的蛋白大多为非典型的 LEA 蛋白,它们与其他分组之间存在着较低的序列同源关系^[6],AkLEA 基因在 NCBI 数据库中进行 Blast,结果显示在与狗枣猕猴桃 LEA 蛋白相比的 98 条 LEA 基因的氨基酸序列中,一致性最高的为 93.15%,最低的为 44.44%。该结果与 Galau 等的研究结



Parasponia andersonii (山黄麻)、*Trema orientale* (东方苔草)、*Quercus suber* (欧洲栓皮栎)、*Nelumbo nucifera* (荷花)、*Actinidia chinensis* var. *Chinensis* (中华猕猴桃原变种)、*Malus domestica* (苹果)、*Cephalotus follicularis* (土瓶草)、*Macleaya cordata* (博落回)、*Theobroma cacao* (可可树)、*Capsicum annuum* (甜椒)、*Capsicum baccatum* (巴卡廷辣椒)、*Solanum lycopersicum* (番茄)、*Helianthus annuus* (向日葵)、*Artemisia annua* (香青蒿)、*Populus trichocarpa* (杨树)、*Punica granatum* (石榴)

图6 蛋白系统进化树分析结果

^[10]相符。LEA2 分组中的非典型 LEA 蛋白,与典型的 LEA 蛋白相比,往往具有较高疏水特性,并且具有较为稳定的结构。典型的 LEA 蛋白具有较强的亲水特性,它们在脱水条件下可以起到亲水性溶质的作用,进而维持蛋白结构的稳定。因此,为了鉴定 *AkLEA* 基因 cDNA 序列是否也具有非典型 LEA 蛋白的这些特性,本研究通过 Prot Param 在线软件进行了疏水性鉴定,但是结果预测该 *AkLEA* 蛋白为亲水性蛋白。这一结果表明了 *AkLEA* 蛋白可能拥有不同于非典型的 LEA 蛋白的功能。选取 17 条氨基酸序列与狗枣猕猴桃的氨基酸序列进行比对,经 MEGA 5.0 分析,结果表明,狗枣猕猴桃 *AkLEA* 与中华猕猴桃变种亲缘关系较近,而与杨树、石榴等植物亲缘关系较远。

参考文献:

- [1] 苏玉彤,张增江,刘长武. 狗枣猕猴桃利用价值及开发前景[J]. 特种经济动植物,2014(6):46-47.
- [2] 李红莉. 东北地区野生猕猴桃生物学特性的分析与研究[J]. 中国林副特产,2013(2):13-15.
- [3] 鲁松,杨楠,熊铁一. LEA 蛋白及其基因研究进展[J]. 四川林业科技,2013(6):26-28,49.
- [4] 王莲哲,王渊,柳超奇,等. 小麦 *TaLEA3-1* 基因的克隆、表达与渗透胁迫功能分析[J]. 分子植物育种,2019,17(14):4524-4528.
- [5] Shao H B, Liang Z S, Shao M A. LEA proteins in higher plants: structure, function, gene expression and regulation[J]. Colloids and Surfaces B - Biointerfaces, 2005, 45(3/4):131-135.
- [6] Shinozaki K, Yamaguchi - Shinozaki K. Molecular responses to dehydration and low temperature; differences and cross-talk between two stress signaling pathways[J]. Current Opinion in Plant Biology, 2000, 3(3):217-223.
- [7] 张林生,赵文明. LEA 蛋白与植物的抗旱性[J]. 植物生理学通讯,2003,39(1):61-66.
- [8] 刘洋,邢鑫,李德全. LEA 蛋白的分类与功能研究进展[J]. 生物技术通报,2011(8):36-43.
- [9] Liu Y, Wang L, Xing X, et al. ZmLEA3, a multifunctional group 3 LEA protein from maize (*Zea mays* L.), is involved in biotic and abiotic stresses[J]. Plant and Cell Physiology, 2013, 54(6):944-959.
- [10] Galau G A, Wang H Y, Hughes D W. Cotton Lea5 and Lea14 encode atypical late embryogenesis - abundant proteins[J]. Plant Physiology, 1993, 101(2):695-696.