

徐婉莉,裴徐梨,荆赞革,等. 辣椒 *actin* 基因电子克隆与生物信息学分析[J]. 江苏农业科学,2014,42(5):46-48.

辣椒 *actin* 基因电子克隆与生物信息学分析

徐婉莉¹,裴徐梨²,荆赞革^{1,2},熊自立¹

(1. 温州科技职业学院,浙江温州 325006; 2. 南京农业大学园艺学院,江苏南京 210095)

摘要:利用拟南芥 *actin* 基因序列为探针,采用电子克隆方法获得辣椒 *actin* 基因 cDNA 序列,对其编码氨基酸序列组成、理化性质、二级结构特点进行分析,并与其他植物 *actin* 的进化关系进行了研究。结果表明,辣椒 *actin* 基因 cDNA 全长为 1 865 bp,包含 1 个完整的开放读码框(1 134 bp),编码 377 个氨基酸,其分子量为 41 700.6 u,等电点为 5.31,为亲水蛋白。辣椒 *actin* 与拟南芥 *actin2* 同源性为 99.0%,进化关系上与番茄 *actin97* 亲缘关系较近。

关键词:辣椒;电子克隆;生物信息学

中图分类号: S641.301 **文献标志码:** A **文章编号:** 1002-1302(2014)05-0046-03

辣椒(*Capsicum annum* L.)别称秦椒、辣子、番椒等,属茄科辣椒属一年生或多年生草本植物,原产于拉丁美洲热带地区。辣椒以其特殊的口感与高维生素 C 含量而深受消费者喜爱,在我国大部分地区均有栽培,是重要的蔬菜作物。随着基因高通量测序技术的迅猛发展,数据库表达序列标签(expressed sequence tags, EST)序列数量激增,电子克隆(in silico cloning)作为快速克隆目标基因技术应用愈加广泛。电子克隆主要基于 EST 数据进行基因分析及试验验证,与传统的克隆方法相比,具有高效率、低成本、操作简单等优点,目前已在多种作物中得到了应用^[1-11]。本研究以拟南芥 *actin* 基因序列为探针,利用电子克隆方法获得 1 条辣椒 *actin* 基因序列,并对该基因及其编码蛋白进行生物信息学分析,旨在为开发辣椒资源提供依据。

1 材料与方法

1.1 材料

以拟南芥 *actin* 基因(GenBank 登录号:U37281)的 cDNA 序列为探针序列,采用电子克隆方法,获得辣椒 *actin* 基因。

1.2 方法

1.2.1 BLAST 搜索辣椒 EST 数据库 以拟南芥 *actin* 基因(GenBank 登录号:U37281)的 cDNA 序列编码区为探针序列,搜索辣椒表达序列标签(EST)数据库。选择 EST 序列,利用 CAP3 在线拼接软件(<http://pbil.univ-lyon1.fr/cap3.php>)进行序列拼接,将拼接后的序列作为探针再次 BLAST 搜索辣椒表达序列标签数据库,直至序列无法延伸为止。最后得到的拼接序列即为电子克隆得到的辣椒 *actin* 基因序列。

1.2.2 辣椒 *actin* 基因序列生物信息学分析 利用 NCBI 开放读码框搜寻软件 ORF finder (<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>)寻找电子克隆到的辣椒 *actin* 基因序列开

放读码。采用多序列比对软件 Clustalw 2.1 将其与拟南芥 *actin2* 基因编码的蛋白进行相似性比对分析。在线分析辣椒 *actin* 编码蛋白的一级结构特点(<http://www.expasy.org/tools/protparam.html>),二级结构在(http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html)进行分析,采用 ProtScale 程序(<http://web.expasy.org/protscale>)分析疏水性。

1.2.3 辣椒 *actin* 基因编码蛋白同源性和进化分析 从 GenBank 中选择多个物种的 *actin* 基因序列,利用 MEGA 5.05 程序对编码氨基酸序列进行同源性比对,利用 Clustal X 软件进行氨基酸序列同源性比对分析,并构建系统发育树。

2 结果与分析

2.1 辣椒 *actin* 基因 EST 序列的搜索、拼接

以辣椒 *actin* 基因编码区为探针,在 NCBI 网站辣椒 EST 数据库中进行检索,得到 7 条(E-value=0)同源性 EST 序列,将序列保存到 FASTA 文件中,采用在线拼接软件 CAP3 进行拼接,得到 3 个 contig,以 contig2 为探针继续对辣椒 EST 数据库进行 BLAST 搜索,得到 21 条序列(E-value=0),拼接得到最终 contig,长度为 1 865 bp,且此 contig 无法再延伸,因此判定其为电子克隆得到的辣椒 *actin* 基因 cDNA 编码序列。

2.2 辣椒 *actin* 基因序列生物信息学分析

2.2.1 辣椒 *actin* 基因序列开放读码分析 利用 ORF finder 软件对辣椒 *actin* 基因的 cDNA 序列进行分析,获得最长开放读码框 1 134 bp,起始密码子位于第 355 bp,终止密码子位于第 1 489 bp,编码 377 个氨基酸(图 1)。

2.2.2 同源分析 利用 Clustalw 1.83 软件对拟南芥与辣椒 *actin* 基因编码的氨基酸序列进行同源比对,结果显示,二者同源性为 99%(图 2)。

2.2.3 辣椒 *actin* 蛋白的一级结构特点 辣椒 *actin* 蛋白的一级结构特点分析结果表明,该蛋白由 Ala、Arg 等 20 种氨基酸组成。其中 Gly 含量最高(8.0%),Trp 最低(1.1%),分子量为 41 700.6 u,等电点为 5.31,分子式为 $C_{1849}H_{2915}N_{493}O_{563}S_{20}$,由 5 840 个原子组成,消光系数为 44 600(280 nm),估计半衰期为 30 h(哺乳动物网状细胞,体外),热稳定指数(II)

收稿日期:2013-08-02

基金项目:浙江省公益性技术应用研究计划(编号:2012C22045)。

作者简介:徐婉莉(1975—),女,硕士,从事蔬菜遗传育种研究。

E-mail:wxw10921@163.com。

通信作者:荆赞革,博士,助理研究员,从事蔬菜遗传育种与生物技术研究。E-mail:jingzange@aliyun.com。

图1 电子克隆的辣椒 *actin* 基因cDNA序列及其推导的氨基酸序列

图2 拟南芥与辣椒*actin*基因编码蛋白的同源性比较

bidopsis thaliana)、番茄 (*Solanum tuberosum*)、水稻 (*Oryza sativa*) 3 个物种 9 条 *actin* 基因编码蛋白序列, 同辣椒 *actin* 蛋白一起采用 MEGA 5.05 软件构建 Neighbor-joining 系统发育树。从图 4 可以看出, 10 条序列可分为两大分支, 番茄 *actin8*

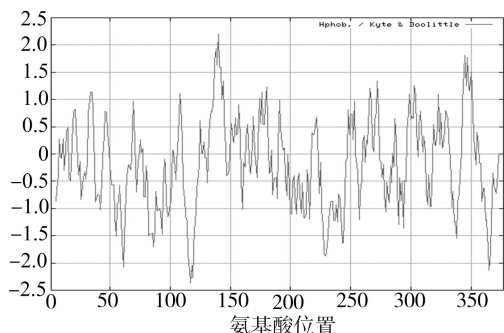


图3 辣椒actin蛋白亲水性分布

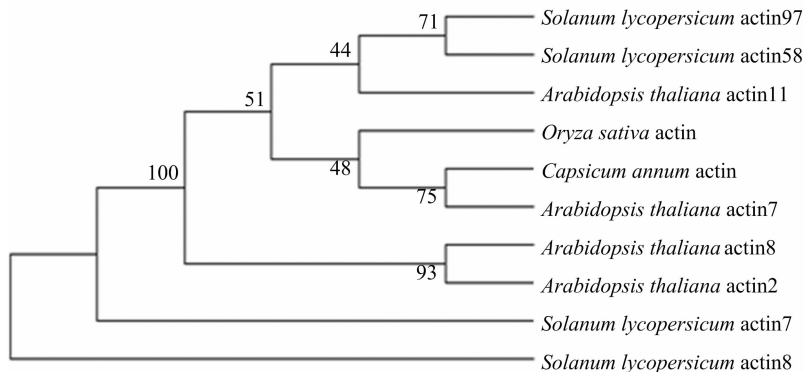


图4 辣椒与部分植物actin蛋白系统进化树

因编码蛋白一般由 375 ~ 377 个氨基酸组成,分子量约 42 ku,是细胞骨架的主要成分之一^[12]。本试验获得的辣椒 *actin* 基因编码 377 个氨基酸,分子量为 41 ku,与多种植物 *actin* 基因核苷酸序列的同源性在 83% 以上,氨基酸同源性在 97% 以上,表明 *actin* 基因的高度保守性,可能与其参与构成细胞骨架的重要功能紧密相关。拟南芥上至少发现了 20 个 *actin* 基因,水稻、杨树、棉花上分别发现了 18 个、22 个、15 个^[13]。不同的基因编码、不同类型的肌动蛋白异形体参与不同的生理活动与生命过程,如细胞器运动、花粉管顶端生长、细胞形状控制、细胞有丝分裂等。*actin* 基因家族因其保守性高,在各种组织中表达相对稳定,被视为看家基因,是研究其他基因表达调控模式的分子内标,在定量与半定量 PCR 研究中用作内参,以确定目标基因的相对表达量^[14]。内参基因 *actin* 的克隆是研究特定植物基因的重要基础,近年来已在多个物种中分离鉴定出 *actin* 基因。

参考文献:

- [1] 马光,梁菲菲,郭继平,等. 白菜黄烷酮-3-羟化酶基因的电子克隆与序列分析[J]. 北方园艺,2012(17):116-118.
- [2] 刘志勇,王孝宣,高建昌,等. 番茄 S-腺苷蛋氨酸脱羧酶基因 *SISAMDC1* 的克隆与序列分析[J]. 园艺学报,2008,35(8):1137-1146.
- [3] 冯爱芹,何晓红,叶绍辉,等. 猪原癌基因 *JunB* 的电子克隆及比较基因组学分析[J]. 江苏农业科学,2012,40(10):33-36.

(XP_004244002)、番茄 *actin7* (XP_004250065) 单组分聚类,番茄 *actin97* (XP_004249286)、番茄 *actin58* (XP_004236747)、拟南芥 *actin11* (NP_187818) 聚为 1 个亚组,电子克隆到的辣椒 *actin* 与番茄 *actin97* (XP_004236747) 亲缘关系最近,聚在 1 个亚组,拟南芥 *actin8* (NP_175350)、拟南芥 *actin2* (NP_188508) 聚为 1 个亚组。

3 结论与讨论

肌动蛋白普遍存在于真核生物中,植物 *actin* 基因起源于共同的祖先,通过多次复制与变异保留下来,因此 *actin* 基因在核苷酸与氨基酸水平上具有高度保守性、同源性。*actin* 基

- [4] 方静平,苏亚春,游倩,等. 甘蔗 β -1,3-葡聚糖酶基因的电子克隆与分析[J]. 生物信息学,2012,10(3):199-207.
- [5] 杨胡,杨军厚. 斑节对虾 1 种 *lectin* 基因的电子克隆及分析[J]. 江苏农业科学,2012,40(8):34-36.
- [6] 马光,郭继平,高小宽,等. 萝卜干细胞决定基因 *WUS* 的电子克隆与序列分析[J]. 生物技术,2012,22(4):46-49.
- [7] 巩元勇,郭书巧,束红梅,等. 陆地棉 *GhNIP5. J* 基因的电子克隆及生物信息学分析[J]. 江苏农业学报,2013,29(3):682-684.
- [8] 畅丽萍,魏琦超,周岩. 马铃薯促分裂原活化蛋白激酶激酶基因的电子克隆与生物信息学分析[J]. 广东农业科学,2012(15):159-162.
- [9] 李萌,张爱华,白雪飞,等. 甜瓜玉米黄质环氧化酶基因的电子克隆及生物信息学分析[J]. 华北农学报,2011,26(增刊1):23-28.
- [10] 王俊生,范小芳,李成伟,等. 小麦 *Ta-UBX1* 基因的电子克隆和生物信息学分析[J]. 江苏农业科学,2012,40(6):29-32.
- [11] 邹晓勇,孙雁霞,何钢,等. 一个玉米 *Mlo* 基因的电子克隆与生物信息学分析[J]. 玉米科学,2011,19(1):148-152.
- [12] 陈颖,王刚,赵俊霞. 高等植物体内的肌动蛋白[J]. 生物学通报,2003,38(1):13-15.
- [13] 杨亚军,王新超,马春雷. 茶树肌动蛋白基因 (*CsActin1*) 全长 cDNA 克隆与生物信息学分析[J]. 植物研究,2012,32(1):69-76.
- [14] 荆赞革,柳李旺,龚义勤,等. 萝卜扩展蛋白基因 *RsEXPB1* 克隆与表达特征分析[J]. 分子植物育种,2009,7(4):801-805.