

粘虫核糖体蛋白 S11 基因 cDNA 的克隆和序列分析

李 柯^{1,2}, 阴 环^{1,2}, 廉振民^{1,3}, 奚耕思¹

(1 陕西师范大学 生命科学学院, 陕西 西安 710062; 2 山西师范大学 生命科学学院, 山西 临汾 041004;

3 延安大学 生命科学学院, 陕西 延安 716000)

【摘要】【目的】克隆和鉴定粘虫核糖体蛋白 S11(Ribosomal Protein S11, RPS11)基因,为昆虫核糖体蛋白功能的研究提供基础资料。【方法】运用 RT-PCR 和 RACE 技术,以粘虫 cDNA 为模板,对 RPS11 基因进行克隆,获得其全长 cDNA 序列,并利用生物信息学方法,对 RPS11 基因全长 cDNA 序列及推测得到的 RPS11 蛋白氨基酸序列进行分析,构建其系统进化树。【结果】获得的粘虫核糖体蛋白 S11 基因(RPS11)cDNA 序列长度为 521 bp,其中包括 26 bp 的 5'非编码区、36 bp 的 3'非编码区和 459 bp 的开放阅读框,编码一个由 152 个氨基酸组成的蛋白,其具有核糖体蛋白 S17 蛋白家族典型特征。推测得到的粘虫 RPS11 蛋白的理论分子质量为 17.737 9 ku,等电点为 10.63,富含 6 种类型的特定功能位点,与同属夜蛾科的烟芽夜蛾(*Heliothis virescens*)的 RPS11 蛋白相似性高达 97%。用 Neighbor-joining(NJ)法构建基于 RPS11 氨基酸序列的系统发育树,结果显示,粘虫 RPS11 与烟芽夜蛾(*H. virescens*)、草地贪夜蛾(*Spodoptera frugiperda*)和家蚕(*Bombyx mori*)等 3 种鳞翅目昆虫的亲缘关系较近。【结论】成功获得了粘虫 RPS11 基因全长序列(GenBank 登录号为 GQ222274),由其编码的核糖体蛋白 RPS11 属于核糖体蛋白 S17 蛋白家族。

【关键词】 粘虫;核糖体蛋白 S11 基因;克隆

【中图分类号】 Q785;Q966

【文献标识码】 A

【文章编号】 1671-9387(2010)04-0171-05

Cloning and sequence analysis of a ribosomal protein S11 gene in the oriental armyworm, *Mythimna separata* (Walker)

LI Ke^{1,2}, YIN Huan^{1,2}, LIAN Zhen-min^{1,3}, XI Geng-si¹

(1 College of Life Sciences, Shaanxi Normal University, Xi'an, Shaanxi 710062, China; 2 College of Life Sciences, Shanxi Normal University, Linfen, Shanxi 041004, China; 3 College of Life Sciences, Yanan University, Yanan, Shaanxi 716000, China)

Abstract: 【Objective】 Ribosomal protein S11(RPS11) gene in *Mythimna separata* (Walker) is cloned and characterized in order to provide scientific data for further investigation on ribosomal proteins function in insects. 【Method】 The full-length cDNA of RPS11 gene is cloned from *M. separata* (Walker) using RT-PCR and RACE technique; the sequence of full-length cDNA and the deduced protein are analyzed using bioinformatics, and the phylogenetic trees of RPS11 is constructed. 【Result】 The full-length cDNA of RPS11 gene in *M. separata* (Walker) is 521 base pairs (bp) and contains a 5'-untranslated region (5'-UTR) of 26 bp and a 3'-untranslated region(3'-UTR) of 36 bp. The open reading frame (ORF) of 459 bp encodes is a 152 amino acid protein which shares typical ribosomal protein S17 family signature and 97% identity with *Heliothis virescens* RPS11. A phylogenetic tree constructed by the neighbor-joining (NJ) method is based on the amino acid sequences of *M. separate* (Walker) RPS11 and other organisms show the relationships among *M. separate* (Walker), *H. virescens*, *Spodoptera frugiperda* and *Bombyx mori* are

* [收稿日期] 2009-10-17

[基金项目] 教育部科学技术研究重点项目(206148);陕西省教育厅科研计划项目(06JK156)

[作者简介] 李 柯(1974-),男,山西陵川人,在读博士,主要从事昆虫生理生态研究。

[通信作者] 廉振民(1955-),男,陕西礼泉人,教授,博士生导师,主要从事昆虫学、生态学及生物多样性研究。

E-mail: lzml169@yau.edu.cn

closer. 【Conclusion】 The full-length cDNA sequence of *RPS11* gene is obtained successfully in *M. separata* (Walker) in this study (GenBank accession no. GQ222274), and the deduced protein based on the cDNA sequence of *RPS11* gene belongs to ribosomal protein S17 family.

Key words: *Mythimna separata* (Walker); *RPS11*; cloning

核糖体是细胞内负责合成蛋白质的细胞器,主要由核糖体 RNA(rRNA)和蛋白质构成。真核细胞核糖体由一个 60 S 大亚基和 40 S 小亚基结合而成。在大亚基中,约有 49 种蛋白质,此外还有 28 S rRNA、5 S rRNA 和 5.8 S rRNA 3 种 rRNA;小亚基大约含 34 种蛋白质,一种 18 S rRNA^[1-2]。核糖体蛋白的命名与蛋白在核糖体的亚基有关,大亚基核糖体蛋白命名为 L1 至 L49,小亚基核糖体蛋白命名为 S1 至 S34^[1]。核糖体蛋白 S11(RPS11)为组成 40 S 小亚基的蛋白成分之一。目前,关于无脊椎动物、脊椎动物、细菌以及植物的 RPS11 基因 cDNA 序列已有不少报道^[3-12],而对 RPS11 功能的研究较少。人类 RPS11 属于核糖体蛋白 S17 蛋白家族,为大肠杆菌 RPS17 直系同源基因。在大肠杆菌中,RPS17 特异地结合在 16 S rRNA 的 5'末端,参与识别终止密码子^[13]。本研究以粘虫(*Mythimna separata* (Walker))作为试验材料,克隆获得粘虫 40 S 核糖体蛋白 S11 基因的 cDNA 序列,并利用生物信息学方法对其进行分析,旨在为昆虫核糖体蛋白功能的研究奠定基础。

1 材料与方 法

1.1 材 料

粘虫 *Mythimna separate* (Walker)由西北农林科技大学无公害农药研究中心提供。培养时室内温度为(23±1)℃,光周期为 L14 : D10(L:Light;D:Dark),相对湿度为 70%~80%。幼虫用小麦苗进行饲养。

1.2 方 法

1.2.1 总 RNA 的提取和 cDNA 第一链的合成

任意选取 5 龄粘虫幼虫 8 头,取其头部立即放入含有液氮的研钵中,充分研磨后,加入 RNAiso Plus (购自 TaKaRa 公司),按照 RNAiso Plus 使用说明书进行总 RNA 的提取。获得总 RNA 后,立即使用 TaKaRa 公司扩增 cDNA 3'末端全长试剂盒(3'-Full RACE Core Set)和 5'末端全长的试剂盒(5'-Full RACE Core Set),在 10 μL 体系中反应合成含 3'末端第一链 cDNA 和 5'末端第一链 cDNA。

1.2.2 引物设计、PCR 扩增及产物的克隆与测序

根据 GenBank 中蛋白 S11 氨基酸序列的保守区域,设计 1 对简并引物 A1 (5'-ACATAGACAARAARTGYCC-3')和 A2 (5'-TGCCCATTCACNGGNAAYGT-3'),将该引物直接用于 3'末端全长 cDNA 的扩增,按 3'-Full RACE Core Set 试剂盒推荐的方法进行。先利用引物 A1 与 3'-RACE outer,以 3'末端全长 cDNA 的第 1 链为模板进行第一轮 PCR 扩增(FIRST-PCR)。扩增反应条件为:94℃预变性 3 min;然后 94℃ 30 s,50℃ 50 s,72℃ 1 min,共 35 个循环;最后 72℃再延伸 10 min。然后用简并引物 A2 与 3'-RACE inner 进行巢式 PCR 扩增(NEST-PCR)。PCR 扩增条件为:94℃预变性 3 min;然后 94℃ 30 s,58℃ 50 s,72℃ 1 min,共 35 个循环;最后 72℃再延伸 10 min。PCR 产物经确认后用 20 g/L 琼脂糖凝胶电泳分离纯化,再克隆至 pMD-19T 载体,转化感受态细胞 *E. coli* DH5α,将 PCR 检测为阳性的克隆,送上海生工生物技术有限公司用 ABI 3730 测序仪进行测序。根据获得的 3'末端核糖体 S11 cDNA 片段序列,设计特异性引物 B1 (5'-CCCTAAAGCACGGCGATAAG-3')和 B2 (5'-TCATCTTCTGGACCACGCCG-3'),用于 5'-RACE cDNA 片段的扩增。FIRST-PCR 反应条件为:94℃预变性 3 min;然后 94℃ 30 s,50℃ 50 s,72℃ 1 min,共 35 个循环;最后 72℃再延伸 10 min。NEST-PCR 条件为:94℃预变性 3 min;然后 94℃ 30 s,65℃ 50 s,72℃ 1 min,共 35 个循环;最后 72℃延伸 10 min。所得 5'-RACE PCR 产物的纯化、克隆和测序方法同 3'-RACE PCR 产物。

1.3 生物信息学分析

采用软件 ProtParam (<http://www.expasy.ch/tools/protparam.html>)分析蛋白质分子质量和等电点;采用 SOPMA 软件(<http://npsa-pbil.ibcp.fr/>)分析蛋白质的二级结构;运用 TMpred 软件(http://www.ch.embnet.org/software/TMPRED_form.html)分析蛋白质跨膜区。采用 MEGA 4.0 软件构建进化树,进化距离分析采用邻接法(Neighbor-joining,NJ),系统树每个分支的统计学显著性分析用bootstrap 进行检验,重复次数为

1 000 次。

2 结果与分析

2.1 粘虫 RPS11 基因 RACE 巢式 PCR 扩增产物的检测

电泳结果表明, 3'-RACE 和 5'-RACE 巢式

PCR 扩增的产物大小分别约为 500, 250 bp(图 1)。将 PCR 产物克隆到 pMD-19T 载体, 然后转化感受态细胞 *E. coli* DH5 α , 再将 PCR 检测为阳性的克隆送上海生工生物技术有限公司进行测序, 结果显示, 3'-RACE 和 5'-RACE 巢式 PCR 扩增的产物长度分别为 491, 251 bp。

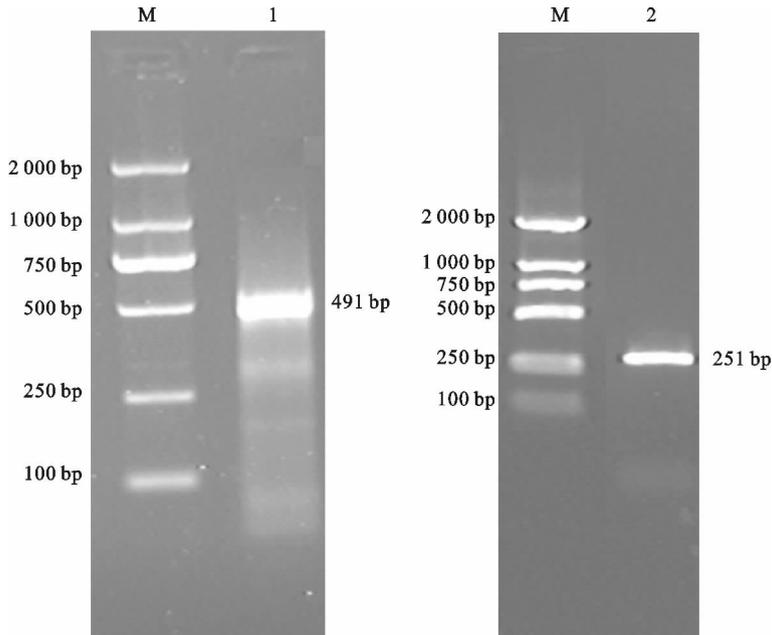


图 1 粘虫 RPS11 基因的 RACE 巢式 PCR 扩增产物

M. DNA marker DL2000; 1. 3'-RACE 巢式 PCR 产物; 2. 5'-RACE 巢式 PCR 产物

Fig. 1 RACE nested PCR products of ribosomal protein S11 (RPS11) gene in *Mythimna separata* (Walker)

M. DNA marker DL2000; 1. Product of 3'-RACE nested PCR; 2. Product of 5'-RACE nested PCR

2.2 粘虫 RPS11 基因全长 cDNA 序列的获得及开放阅读框的识别

将经过扩增、测序得到的 3'端 cDNA 片段与 5'端 cDNA 片段序列, 通过 ContigExpress 软件拼接, 得到粘虫 RPS11 基因全长 cDNA 序列为 521 bp。采用 NCBI(The National Center for Biotechnology Information) 上的在线工具 ORF Finder (<http://www.ncbi.nlm.nih.gov/gorf/orfig.cgi>) 进行开放阅读框(Open Reading Frame, ORF)识别, 结果见图 2。图 2 显示, 起始密码子 ATG 位于第 27~29 位核苷酸, 终止密码子 TAA 位于第 482~485 位核苷酸, 开放阅读框全长为 459 bp, 编码 152 个氨基酸。5'非翻译区(5'-untranslated region, 5'-UTR) 的长度为 26 bp, 3'非翻译区(3'-untranslated region, 3'-UTR) 的长度为 36 bp, 其中在 3'-UTR 有典型的加尾信号 AATAAA, 可见长度为 18 bp 的 PloyA 尾。采用 NCBI 上的在线工具 BLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) 将翻译所得的蛋白质

序列进行同源性搜索, 结果发现, 与同属夜蛾科烟芽夜蛾(*Heliothis virescens*) 的 RPS11 蛋白相似性高达 97%。由此可判定该序列为粘虫 RPS11 基因的全长序列, 编码一个完整的 RPS11 蛋白质序列, 基因序列已提交 GenBank, 登录号为 GQ222274。

2.3 粘虫 RPS11 蛋白质的结构和功能分析

采用在线工具 ProtParam (<http://www.expasy.ch/tools/protparam.html>), 对由 cDNA 序列推测的粘虫 RPS11 蛋白质序列分析后发现, 其理论分子质量为 17. 737 9 ku, 等电点为 10. 63。用 Proscan 软件 (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_prosite.html) 进行分析, 结果显示, 该蛋白富含 6 种类型的特定功能位点(图 2), 分别为 2 个 N-糖基化位点(59~62: NV-SI, 102~105: NMSV)、6 个蛋白激酶 C 磷酸化位点(5~7: TER, 40~42: TPR, 61~63: SIR, 131~133: TIR, 145~147: SKK, 148~150: SFK)、3 个酪氨酸蛋白激酶 II 磷酸化位点(40~43: TPRES, 48~51: TYID,

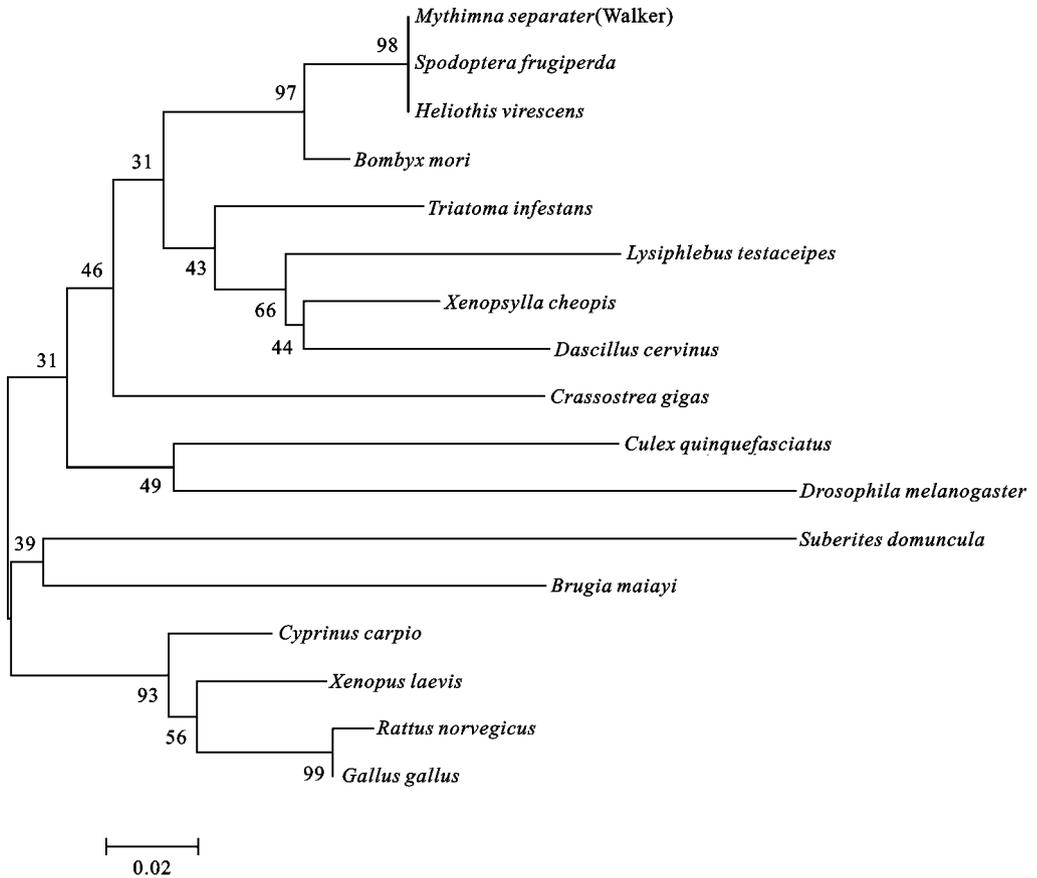


图 3 基于 RPS11 氨基酸序列的粘虫和其他 16 种物种的系统发育树

Fig. 3 Phylogenetic tree of RPS11 from *Mythimna separata* (Walker) and 16 other species

3 讨 论

本研究运用 RT-PCR 和 RACE 技术,克隆获得了 521 bp 的粘虫 RPS11 基因全长 cDNA 序列,其中包括 26 bp 的 5'-UTR、36 bp 的 3'-UTR 和 459 bp 的开放阅读框。基于 RPS11 基因的 cDNA 序列推测的蛋白质由 152 个氨基酸组成。NCBI 的 Blast P 同源性检索结果显示,粘虫 RPS11 属于核糖体蛋白 S17 蛋白家族。

由本研究根据 RPS11 氨基酸序列构建的系统发育树可知,马来丝虫 (*B. malayi*) 和皮海绵 (*S. domuncula*) 与脊椎动物聚为一支,太平洋牡蛎 (*C. gigas*) 与昆虫聚为一类,这表明 RPS11 蛋白序列不适合作为门以上分类阶元的分子标记,这与 Müller 等^[14]的结论一致。本试验发现,尽管在系统发育树中门以下阶元物种的聚类效果较好,但所采用的蛋白序列数量有限,RPS11 蛋白序列是否能作为门以下分类阶元有效的分子标记,还需以大量物种的 RPS11 蛋白序列进一步证实。许多研究表明,核糖体蛋白除参与核糖体构成和蛋白质合成外,还参与

复制、转录、RNA 加工、DNA 修复等过程,并作用于细胞的增殖、凋亡、发育等多种调控及恶性转化过程^[15],但关于 RPS17 蛋白家族的功能报道还较少。本试验首次成功获得了粘虫核糖体蛋白 S11 基因的全长 cDNA,是对昆虫核糖体蛋白研究的重要补充,为今后更多核糖体蛋白序列的获得及相关功能研究提供了基础资料。

志谢:感谢西北农林科技大学无公害农药研究中心张兴教授对粘虫虫源的惠赠!

[参考文献]

- [1] Alberts B, Johnson A, Lewis J, et al. The molecular biology of the cell [M]. 4th edition. New York and London: Garland Publishing, 2002: 342.
- [2] Doudna J A, Rath V L. Structure and function of the eukaryotic ribosome: the next frontier [J]. Cell, 2002, 109(2): 153-156.
- [3] Annesi F, Vespignani I, Amaldi F, et al. *Xenopus laevis* ribosomal-protein S11: cloning and sequencing of the cDNA and primary structure of the protein [J]. Biochemical and Biophysical Research Communications, 1994, 203(2): 768-772.

(下转第 182 页)