

棉铃虫泛素基因的克隆及序列分析^{*}

李朝飞, 于航, 潘丽晶, 杨凯, 庞义

(中山大学生物防治国家重点实验室//昆虫学研究所, 广东 广州 510275)

摘要: 泛素介导的泛素-蛋白酶体通路 (Ubiquitin proteasome pathway, UPP) 是真核细胞内依赖 ATP 的非溶酶体蛋白质降解途径, 该途径对细胞内蛋白的选择性降解起着重要作用。设计一对简并引物, 从棉铃虫 *Helicoverpa armigera* 卵巢细胞 Ha831 中克隆了泛素基因的编码区, GenBank 登录号 AY456195。序列分析表明, 该编码区的长度为 228 bp, 编码 76 个氨基酸, 其编码蛋白的相对分子质量为 8 560, 等电点为 6.56。同源性比较发现, 棉铃虫泛素基因与其他真核生物泛素基因在氨基酸水平上具有 96% 以上的相似性, 而与棉铃虫核多角体病毒泛素的相似性为 76%, 所有已知的泛素关键功能位点在该泛素蛋白中均保守存在。

关键词: 棉铃虫 *Helicoverpa armigera*; 泛素; 序列分析

中图分类号: Q965 **文献标识码:** A **文章编号:** 0529-6579 (2005) 01-0061-04

泛素 (ubiquitin) 由 76 个氨基酸组成, 相对分子质量约为 8 600, 是真核生物内高度保守的一种多肽。从不同种属的真核生物得到的泛素的一级结构几乎相同, 对植物、酵母、动物的泛素的分析表明, 它们的一级结构仅有 1—3 个氨基酸残基的不同。目前所知, 泛素主要通过 ATP 依赖的泛素-蛋白酶体通路 (ubiquitin proteasome pathway, UPP) 高效并高度选择性地对胞质、胞核内蛋白进行降解, 尤其是短命的功能蛋白、癌基因产物和变性、变构蛋白等, 应急状态下也降解细胞内其它结构蛋白。UPP 已被认为是十分重要的细胞功能调节因素, 该蛋白水解通路的发现被认为是细胞周期调控、细胞恶性转化和免疫学研究的重大转折。此外, 泛素作为分子伴侣参与核糖体的生物发生、细胞内吞及其它多种生理功能^[1]。除了真核生物以外, 昆虫病毒是目前发现惟一编码泛素的病毒。苜蓿丫纹夜蛾核多角体病毒 (*Autographa californica* nucleopolyhedrovirus, AcMNPV) 泛素基因的缺失并不影响病毒的复制, 但芽生型病毒粒子的滴度明显降低^[2]。

棉铃虫单粒核多角体病毒 (*Helicoverpa armigera* single nucleocapsid nucleopolyhedrovirus, HaSNPV) 的基因组中也存在一个泛素基因^[3]。为了了解 HaSNPV 侵染过程中, 病毒或病毒泛素与宿主泛素-蛋白酶体通路 (UPP) 的相互作用机理, 本文设计一对简并引物, 从棉铃虫 *Helicoverpa armigera* 卵

巢细胞中克隆了泛素基因, 这为研究泛素在病毒感染中的作用奠定了基础。

1 材料与方法

1.1 材料

1.1.1 细胞, 菌株与质粒 棉铃虫 *Helicoverpa armigera* 卵巢细胞 Ha831 与大肠杆菌菌株 TG1 由本实验室保存, 克隆载体 pMD18-T 购自 TaKaRa 公司。

1.1.2 主要试剂 限制性内切酶、核酸相对分子质量标准 and Taq DNA 聚合酶均购自 TaKaRa 公司。其他试剂均为国产分析纯。

1.2 方法

1.2.1 简并引物设计与合成 根据已知人、大麦、酵母、果蝇及鳞翅目昆虫等泛素蛋白氨基酸序列的高度保守性, 设计一对简并引物。

正向引物:

5'-ATGC A(AG) AT(CT) TT(CT) GTT AA-3',

氨基酸序列: MQI FVKLT;

反向引物:

5'-(AG) CC ACC TCG(CGA) AGIC(TG) (CGA) A(GA) (CG) AC-3',

氨基酸序列: LRLRGG。

引物由上海博亚公司合成。

1.2.2 总 RNA 的提取 总 RNA 的提取, 参考文献 [4] 进行。将适量的 Ha831 细胞在液氮冷冻条件下充分研磨, 依次向研磨的细胞中加入 500 μ L 异硫

* 收稿日期: 2004-02-23

基金项目: 国家重点基础规划“973”资助项目 (G2000016209); 中山大学青年教师启动基金资助项目

作者简介: 李朝飞 (1973 年生), 男, 博士; 通讯联系人: 庞义; E-mail: ls12@zsu.edu.cn

氰酸胍变性溶液、50 μ L 乙酸钠（2 mol/L，pH 4）、500 μ L 水饱和酚和 100 μ L 氯仿，混匀，冰上静置 15 min；4 $^{\circ}$ C，12 000 g 离心 15 min；将水相转移至另一新的 1.5 mL 微量离心管中，加入 500 μ L 异丙醇后置于-70 $^{\circ}$ C 30 min；12 000 g 离心 10 min，用 200 μ L 变性溶液重悬沉淀，加入 200 μ L 异丙醇和 60 μ L 乙酸钠（2 mol/L，pH4），混匀，置于-70 $^{\circ}$ C 30 min；12 000 g 离心 20 min，用 φ =75%乙醇淋洗 RNA，然后离心 5 min；晾干 RNA 沉淀，并用适量的无 RNA 酶水溶解。

1.2.3 第一链 cDNA 的合成 在新的 0.5 mL Eppendorf 管中，依次加入 2 μ L 总 RNA（ $\sim$ 2 μ g），1 μ L OligodT15（50 pmol/L），0.5 μ L RNasin（40 U/ μ L），1 μ L AMV reverse transcriptase（10 U/ μ L，Promega），2 μ L dNTPs（25 mmol/L/each）；5 $\times$ RT buffer 4 μ L，9.5 μ L 无 RNA 酶水。在 PCR 仪中完成反转录：42 $^{\circ}$ C，30 min；95 $^{\circ}$ C，5 min；冰浴 2 min。

1.2.4 PCR 扩增 PCR 反应条件为：50 μ L 反应体系中含 5 μ L 10 $\times$ PCR buffer，0.01 μ mole/each dNTP，50 pmol/each primer，3 units Taq Polymerase， $\sim$ 3 ng cDNA。样品先 94 $^{\circ}$ C 预变性 3 min，然后按下述条件进行 PCR 扩增反应：94 $^{\circ}$ C、30 s，45 $^{\circ}$ C、1 min，72 $^{\circ}$ C、1 min，30 个循环后 72 $^{\circ}$ C、10 min，扩增完毕后置 4 $^{\circ}$ C 终止反应。

1.2.5 PCR 产物的克隆与测序 参照文献[5]，PCR 产物经琼脂糖凝胶电泳，选择大小与所设计的棉铃虫泛素基因片段相符合的单一 DNA 条带，回收、纯化；纯化的 PCR 产物与 pMD18-T 载体连接，重组子 pHaUbi 转化大肠杆菌 TG1；挑取单个白色菌落提取质粒，酶切鉴定后，送交上海博亚公司进行测序。

1.2.6 序列分析 应用 DNAsis 分析软件将所得 DNA 序列翻译成蛋白质序列，并用 BLAST 搜索引擎查询 GenBank/EMBL 数据库，获得同源的核酸与蛋白质序列。多序列比较用 CLUSTAL W（1.82）^[9] 与 Genedoc 软件进行。

2 结 果

2.1 PCR 扩增产物及重组质粒酶切鉴定

琼脂糖凝胶电泳检测结果表明，PCR 扩增产物为约 230 bp 的单一片段（图 1，泳道 2），而挑选的阳性克隆，经质粒抽提，EcoRI、PstI 双酶切，则出现 2 条 DNA 条带，位置与预测一致（图 1，泳道 3）。

2.2 棉铃虫泛素基因的编码区

根据真核生物泛素蛋白 N-端和 C-端高度保



图 1 PCR 产物与重组质粒酶切鉴定
Fig 1 Identification of the recombinant plasmid
(restriction analysis using *EcoRI* / *PstI*)
and the PCR product

- 1: DL 2000 DNA 相对分子质量标准;
- 2: 泛素 PCR 产物;
- 3: pHaUbi 重组质粒酶切产物 (*EcoRI* / *PstI*);
- 4: 1 kb DNA 相对分子质量标准

守的氨基酸序列 MQIFVKTLT 和 LRLRCG，设计一对简并引物扩增棉铃虫泛素编码区。PCR 产物测序结果表明，棉铃虫泛素基因开放阅读框长度为 228 bp，编码 76 个氨基酸（图 2），预计编码蛋白的相对分子质量为 8 560，等电点为 6.56。

1	ATG CAG ATC TTT GTG AAA ACA CTA ACT GGT	30
1	M Q I F V K T L T G	10
31	AAG ACC ATC ACC CTT GAG GTC GAG CCT TCT	60
11	K T I T L E V E P S	20
61	GAC ACG ATC GAA AAT GTT AAA GCC AAA ATC	90
21	D T I E N V K A K I	30
91	CAG GAT AAG GAG GGT ATC CCA CCT GAC CAG	120
31	Q D K E G I P P D Q	40
121	CAG CGA TTG ATC TTC GCC GGA AAG CAG TTG	150
41	Q R L I F A G K Q L	50
151	GAG GAT GGG CGG ACC TTG TCT GAT TAC AAT	180
51	E D G R T L S D Y N	60
181	ATT CAG AAA GAA TCC ACT CTC CAC TTG GTG	210
61	I Q K E S T L H L V	70
211	CTG CGC CTG CGA CGT GGC	228
71	L R L R G G	76

图 2 棉铃虫泛素基因核苷酸序列
及其推衍的氨基酸序列

Fig 2 Nucleotide and deduced amino acid sequence of
the coding sequence of *Helicoverpa armigera* ubiquitin gene

2.3 棉铃虫泛素与其他物种泛素的序列比较

运用 Clustal W 软件对棉铃虫泛素基因核苷酸序列及氨基酸序列与其他真核生物泛素的比较结果见图 3、4。由图中可以看出，棉铃虫泛素基因的核苷酸及氨基酸序列与其他生物泛素极其相似，其

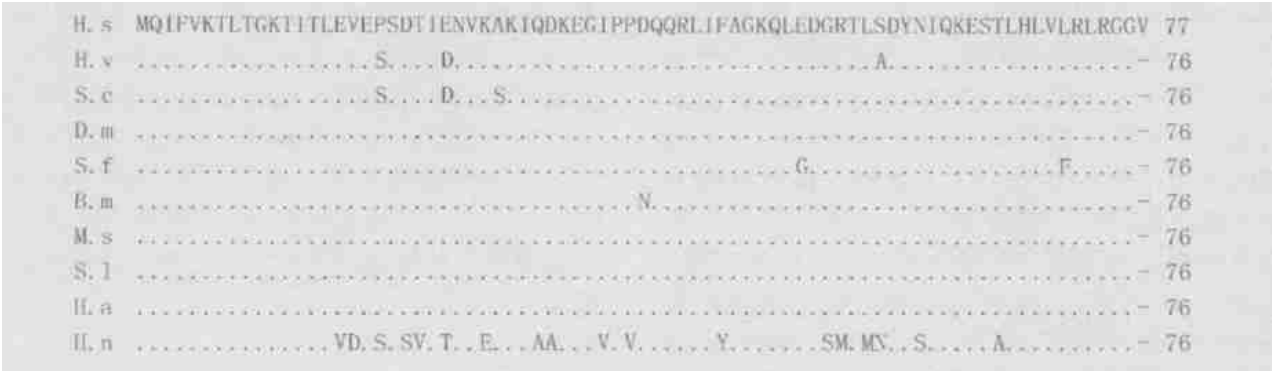


图 3 泛素蛋白氨基酸序列比较

Fig 3 Alignment of the amino acid sequence of ubiquitin proteins

Homo sapiens (H. s) (GenBank accession no. M78113); *Hordeum vulgare* (H. v) (No. X04133); *Saccharomyces cerevisiae* (S. c) (No. X05731); *Drosophila melanogaster* (D. m) (No. X59943); *Spodoptera frugiperda* (S. f) (No. M30306); *Bombyx mori* (B. m) (No. AF308163); *Manduca sexta* (M. s) (No. U61404); *Spodoptera litura* (S. l) (No. AY436066); *Helicoverpa armigera* (H. a); *Helicoverpa armigera* single nucleopolyhedrovirus (H. n)^[3]

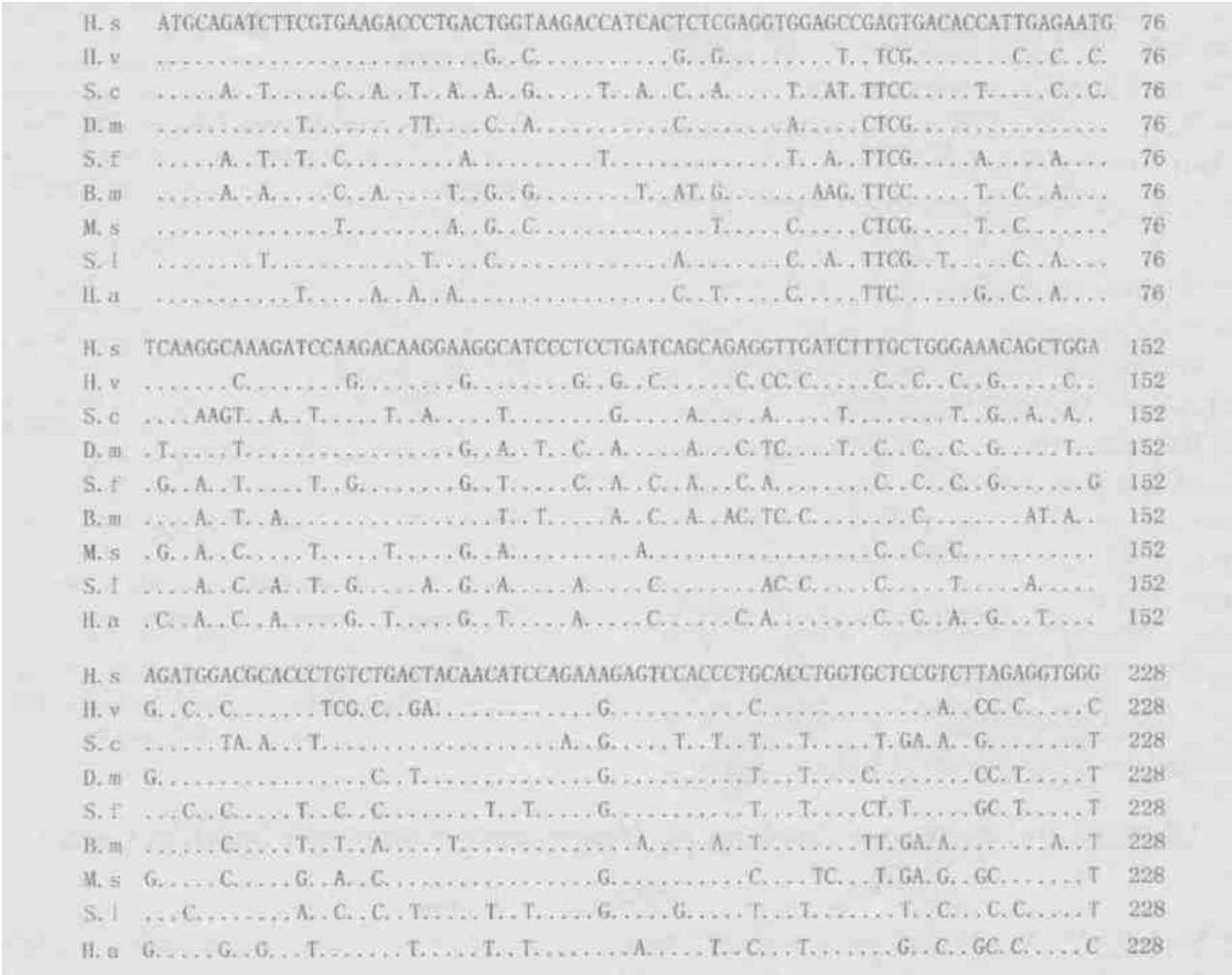


图 4 泛素基因开放阅读框 (ORF) 核苷酸序列比较

Fig 4 Alignment of the nucleotide sequence of ubiquitins' ORF

Homo sapiens (H. s) (GenBank accession no. M78113); *Hordeum vulgare* (H. v) (No. X04133); *Saccharomyces cerevisiae* (S. c) (No. X05731); *Drosophila melanogaster* (D. m) (No. X59943); *Spodoptera frugiperda* (S. f) (No. M30306); *Bombyx mori* (B. m) (No. AF308163); *Manduca sexta* (M. s) (No. U61404); *Spodoptera litura* (S. l) (No. AY436066); *Helicoverpa armigera* (H. a); *Helicoverpa armigera* single nucleopolyhedrovirus (H. n)^[3]

核苷酸序列与其他真核生物泛素的相似性为 74%~77%。大多数碱基差异位点出现在密码子的摆动位置上（图 4）。相比较核苷酸序列，棉铃虫泛素与其他真核生物泛素的氨基酸序列仅有 1~3 个氨基酸的不同（图 3），它与人、果蝇、斜纹夜蛾及烟草天蛾泛素的同源性和为 100%；与家蚕泛素的同源性和为 98%；与草地贪夜蛾泛素的同源性和为 97%；与大麦、酿酒酵母泛素的同源性和为 96%；与 *HaSNPV* 泛素的同源性和为 76%。所有已知与泛素功能相关的位点在棉铃虫泛素蛋白序列中均存在。

3 讨论

自从 20 世纪 80 年代初以来，泛素及泛素—蛋白酶体通路一直是生物学领域的研究热点之一。近年来，大量的研究结果表明，泛素除了与靶蛋白共价结合参与细胞内蛋白的选择性降解外，还参与细胞周期调控、DNA 修复、信号传导、转录调控、细胞凋亡以及机体免疫等多种生命活动过程^[1]。最近，有研究结果表明，家蚕核多角体病毒 [*Bombyx mori* (Bm) NPV] 编码的 IAP2、IE2 和 PE38 3 个指环蛋白具有泛素—蛋白酶体通路中 E3 连接酶的功能^[7]。而且，从 20 世纪 80 年代末以来，人们先后在一些植物病毒和动物病毒的病毒结构中发现了泛素或泛素化的病毒蛋白^[2]。最近，大量的研究结果表明，RNA 病毒利用宿主细胞的泛素或泛素介导的蛋白质降解系统逃避免疫系统的监控、促进病毒复制以及进行病毒粒子的组装和释放^[8]。目前，国内外仅报道了草地贪夜蛾 (GenBank Accession No. M30306)、烟草天蛾 (No. U61404)、果蝇 (No. X59943)、家蚕 (No. AF308163) 及斜纹夜蛾 (No. AY436066) 等 5 种昆虫的泛素基因序列。研究发现，泛素基因在昆虫的中肠、脂肪体、马氏管、飞行肌等组织中具有高水平的表达^[9-10]。迄今没有涉及病理条件下昆虫病毒侵染过程中与宿主细胞泛素—蛋白酶体通路相互作用机理的研究报道。为了了解

泛素与昆虫病毒侵染的相互关系，本文根据已知真核生物泛素基因序列，设计一对简并引物从棉铃虫卵巢细胞中扩增和克隆了泛素基因。序列分析表明，棉铃虫泛素与其他真核生物泛素的氨基酸序列仅有 1~3 个氨基酸的不同，所有已知与泛素功能相关的位点在棉铃虫泛素序列中均保守存在。关于该基因的功能及其与 *HaSNPV* 侵染的相互关系在进一步研究中。

参考文献:

[1] GLICKMAN M H, CIECHANOVER A. The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction [J]. *Physiol Rev*, 2002, 82: 373-428.

[2] 李朝飞, 庞义. 泛素—蛋白水解酶复合体通路与病毒侵染[J]. *生物工程学报*, 2004, 22(2): 1-7.

[3] CHEN X, IJEL W F, TARCHINI R, et al. The sequence of the *Helicoverpa armigera* single nucleocapsid nucleopolyhedrovirus genome [J]. *J Gen Virol*, 2001, 82(1): 241-257.

[4] CHOMCZYNSKI P, SACCHI N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction [J]. *Anal Biochem*, 1987, 162: 156-159.

[5] SAMBROOK J T, MANIATS E F. Molecular cloning: A laboratory manual [M]. New York: Cold Spring Harbor Laboratory, 1982.

[6] THOMPSON J D, HIGGINS D G, GIBSON T J. Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice [J]. *Nucleic Acids Res*, 1994, 22: 4673-4680.

[7] IMAI N, MATSUDA N, TANAKA K, et al. Ubiquitin ligase activities of *Bombyx mori* nucleopolyhedrovirus RING finger proteins [J]. *J Virol*, 2003, 77(2): 923-930.

[8] VOLKER M V. Ubiquitin in retrovirus assembly: actor or bystander [J]. *Proc Natl Acad Sci USA*, 2000, 97(24): 12945-12947.

[9] BARRIO R A, DEL ARCO H L, ARRIBAS C. Structure and expression of the *Drosophila* ubiquitin-80 amino acid fusion protein gene [J]. *Biochem J*, 1994, 302: 237-244.

[10] BISHOFF S T, SCHWARTZ L M. Characterization of a ubiquitin fusion gene from the tobacco hawk moth, *Manduca sexta* [J]. *Nucleic Acids Res*, 1990, 18: 6039-6043.

Cloning and Sequence Analysis of *Helicoverpa armigera* Ubiquitin Gene

LI Zhao fei, YU Hang, PAN Li jin, YANG Kai, PANG Yi

(State Key Laboratory for Biocontrol and Institute of Entomology, Sun Yat sen University, Guangzhou 510275, China)

Abstract: Ubiquitin plays a very important role in regulated non-lysosomal ATP dependent protein degradation. The coding sequence of *Helicoverpa armigera* ubiquitin gene was isolated (GenBank Accession No. AY456195). The length of this ORF is 228 bp, encoding a protein with Mr of 8 560 and isoelectric point of 6. 56. Multiple sequence alignment indicated that *H. armigera* ubiquitin is very similar to the homologous proteins of other eukaryotic species and it has more than 96% identity with other eukaryotic ubiquitins and 76% with *Helicoverpa armigera* single nucleopolyhedrovirus (*HaSPV*) ubiquitin at amino acid level. All known key functional sites are still existed in *H. armigera* ubiquitin.

Key words: *Helicoverpa armigera*; ubiquitin; sequence analysis