

棉铃虫核型多角体病毒 ORF44 基因的 序列分析及原核表达

李 萌¹,张雅林¹,王 敦^{1,2}

(1 西北农林科技大学 植保资源与病虫害治理教育部重点实验室,陕西 杨凌 712100; 2. 病毒学国家重点实验室,湖北 武汉 430071)

[摘 要] 棉铃虫核型多角体病毒 (HearNPV) ORF44 是一个功能未知的基因,研究对 HearNPV ORF44 基因的序列及原核表达进行了初步研究。结果表明,HearNPV ORF44 基因全长 1 137 nt,编码 378aa,预测编码蛋白分子量为 42.7 ku。序列分析显示,ORF44 具有双相启动子特征序列,同时兼有杆状病毒早期和晚期基因转录启动子特征序列。ORF44 的推导氨基酸序列中含有 1 个强跨膜螺旋区域,32 个磷酸化位点和 4 个糖基化位点。除 HzNPV ORF45 外,未发现与 ORF44 氨基酸同源性高于 30% 的基因。构建了 pET-44 表达载体,并在大肠杆菌 BL21 中得到了高效表达,表达的融合蛋白分子量为 44 ku,与预测编码蛋白分子量大小相符。序列分析结果表明,ORF44 是 HearNPV 的一个特有基因,ORF44 融合蛋白的成功表达为深入研究 HearNPV ORF44 奠定了基础。

[关键词] 棉铃虫核型多角体病毒 (HearNPV); ORF44 基因; 序列分析; 原核表达

[中图分类号] Q786

[文献标识码] A

[文章编号] 1671-9387(2007)06-0130-05

Sequence analysis and prokaryotic expression of HearNPV ORF44

LI Meng¹, ZHANG Ya-lin¹, WANG Dun^{1,2}

(1 Key Laboratory of Plant Protection Resources and Pest Management, National Ministry of Education, Entomological Museum, Northwest A & F University, Yangling, Shaanxi 712100, China; 2 State Key Laboratory of Virology, Wuhan, Hubei 430071, China)

Abstract : The function of HearNPV ORF44 is unknown for the present, thus the HearNPV ORF44 was studied and the primary results were reported in this article. HearNPV ORF44 is 1 137 nt long, encoding a putative protein of 378aa with a predicted size of 42.7 ku. Both baculovirus consensus early and late promoter motifs are found in ORF44. One transmembrane helices region, 32 putative phosphorylation sites and 4 N-linked glycosylation sites are found in the predicted ORF44 protein. No homologue that shares an identity over 30% through NPVs and GV is found in GenBank and SWISS-PROT. A pET-44 plasmid is constructed and ORF44 is expressed in the *Escherichia coli* (BL21). The molecular weight of expressed fusion protein is 44 ku, which conforms to the predicted size. The results show that ORF44 is unique to HearNPV among NPVs, and the expression of fused ORF44 protein provides the basic information on the analysis of HearNPV ORF44 function.

Key words : HearNPV; ORF44; sequencing analysis; prokaryotic expression

棉铃虫 (*Helioverpa armigera*) 是我国重要的农业害虫之一,常给农业生产带来巨大的经济损失。由于化学农药的长期大量使用,棉铃虫的抗药性剧增,而且污染了环境,使棉铃虫的可持续治理面

临着重大挑战。棉铃虫核型多角体病毒 (*Helioverpa armigera* nucleopolyhedrovirus, HearNPV) 是棉铃虫专一性病原微生物,也是我国第一个注册使用的生物农药^[1],它具有寄主范围特异、对靶标高毒

†收稿日期] 2006-10-30

[基金项目] 病毒学国家重点实验室开放课题(-2005001); 西北农林科技大学校长基金(2005-02)

[作者简介] 李 萌(1982-),男,河南长葛人,在读硕士,主要从事病毒分子生物学研究。

[通讯作者] 王 敦(1973-),男,青海西宁人,副教授,博士,硕士生导师,主要从事昆虫和病毒分子生物学研究。

力、可引起区域性昆虫流行病、无药害、对环境稳定、使用安全、害虫不产生抗药性等优点,是理想的害虫控制因子。然而,传统病毒杀虫剂的杀虫作用缓慢,杀虫范围窄,极大地限制了其推广使用。因此,对 HearNPV 基因进行深入研究,不但有助于更清楚了解其分子生物学背景,而且对病毒杀虫剂的遗传改良和新的杆状病毒-昆虫细胞表达系统的建立具有重要意义。

张传溪等^[2-3]测定了 HearNPV C1 株基因组全序列(NC_003094),构建了基因组详细物理图谱,定位了一些主要基因。到目前为止,HearNPV 的一些基因已见研究报道,如 polyhedrin^[4]、ORF29^[5]、ORF33^[6]、ORF128^[7]、ORF135^[8]以及 ORF83^[9],但在 HearNPV C1 株和其他已报道的杆状病毒的同源基因中,ORF44 尚无研究报道。本研究通过对 HearNPV ORF44 基因的序列分析、原核表达进行研究,以探讨该基因的特性,旨在为 HearNPV 功能基因组学研究提供基础数据。

1 材料与方法

1.1 材料

HearNPV C1 病毒株,由浙江大学应用昆虫研究所张传溪教授惠赠;大肠杆菌 TG1、BL21,载体 pET-28a(+),均由西北农林科技大学植保资源与病虫害治理教育部重点实验室保存。

pGEM-T-easy vector、T4 DNA ligase、Taq DNA 聚合酶及相应 Buffer,均为 Promega 公司产品;限制性内切酶 *Bam*H、*Xho* 及其配套 Buffer、dNTPs、DL2000 DNA Marker,均为 Takara 公司产品;Gel Extraction Kit 凝胶回收试剂盒,为安徽优晶生物工程有限公司产品;辣根过氧化物酶 HRP 标记的 Goat Anti Mouse IgG,购自华美生物工程公司;Anti-His antibody(purified from Mouse serum)购自 Pharmacia 公司;RNase、硝酸纤维素膜(NC)、 β -巯基乙醇,均购自上海 Sangon 公司;琼脂糖 Agarose(ULTRA PURE),为 Amresco 进口分装产品;其他试剂均为国产分析纯。

1.2 方法

1.2.1 序列分析 分析所用的数据通过 BLASTP 程序从 GenBank/EMBL 和 SWISS-PROT 获得;由基因序列推测氨基酸序列用 DNAMAN(Ver 5.2.0)(Lynnon Biosoft, USA)进行编辑分析;功能域、信号肽序列、后修饰等的预测,应用 ScanProsite、Motifscan、SignalP、TmPred、NetNGlyc 和 Net-

Phos 2.0 Server 软件上网分析(<http://ca.expasy.org>)。

1.2.2 原核表达 PCR 反应。根据已发表的 HearNPV C1 株基因组序列设计引物^[2],引物由上海 Sangon 公司合成,其中正向引物含有 *Bam*H 酶切位点,反向引物含有 *Xho* 酶切位点。

ORF44F:5' AGGATCCATGAGCAATCCCA GC;

ORF44R:5' A CTCGAGTTAATAGCGAAACG。

反应体系(25 μ L):模板 DNA 0.5 μ L,正、反向引物各 0.5 μ L,10 $\times$ Buffer(不含 Mg^{2+})2.5 μ L, $MgCl_2$ (25 mmol/L)2 μ L, dNTPs 0.5 μ L, Taq DNA 聚合酶 0.3 μ L,灭菌双蒸水 18.5 μ L。

反应程序:94 预变性 3 min;94 变性 1 min,55 退火 1 min,72 延伸 1.5 min,30 个循环;最后 72 延伸 20 min,产物于 4 保存。

表达载体的构建。用 Gel Extraction Kit 凝胶回收试剂盒从 1.2%琼脂糖凝胶纯化 ORF44 PCR 产物,纯化片段与 pGEM-T-easy vector 连接,重组载体记为 pGT-44,将 pGT-44 转化 TG1 感受态细胞,蓝白斑筛选,质粒 PCR 和 *Bam*H、*Xho* 双酶切鉴定,鉴定无误后送交上海奇骏生物公司测序。

将鉴定无误的 pGT-44 和 pET-28a(+)质粒 *Bam*H、*Xho* 双酶切,Gel Extraction Kit 凝胶回收试剂盒从 1.2%琼脂糖凝胶回收 ORF44 片段和线性 pET-28a(+)质粒,连接 ORF44 与 pET-28a(+),重组载体记为 pET-44,将 pET-44 转化 TG1 感受态细胞,卡那抗性筛选,质粒 PCR 和 *Bam*H、*Xho* 双酶切鉴定。

ORF44 在大肠杆菌中的表达。小量抽提 pET-44,将 pET-44 转化表达菌株 BL21 感受态细胞,鉴定无误后以终浓度 1.5 mmol/L 的 IPTG 在 28 诱导表达 5 h,表达产物进行 12%聚丙烯酰胺凝胶电泳分析。

碱法小量抽提质粒, $CaCl_2$ 法感受态细胞制备、转化及聚丙烯酰胺凝胶电泳的操作方法均参考文献[10],DNA 片段回收和酶切反应的具体操作按厂家说明书进行。

Western blot 分析。Western blot 分析采用一抗为 6 $\times$ His 融合蛋白单抗(Anti-His antibody),二抗为辣根过氧化物酶 HRP 标记的 Goat Anti Mouse IgG。蛋白转移采用 BIO-RAD 半干转膜仪在 15 V 电转 30 min,其他步骤参见文献[11]。

2 结果与分析

2.1 HearNPV ORF44 基因序列分析

HearNPV ORF44 基因位于基因组 38 484 ~

39 620 bp ,基因全长 1 137 nt ,转录方向与棉铃虫核型多角体病毒多角体蛋白基因方向相反 ,编码 378aa ,预测编码蛋白分子量为 42. 7 ku ,pI 值为 9.04(图 1)。

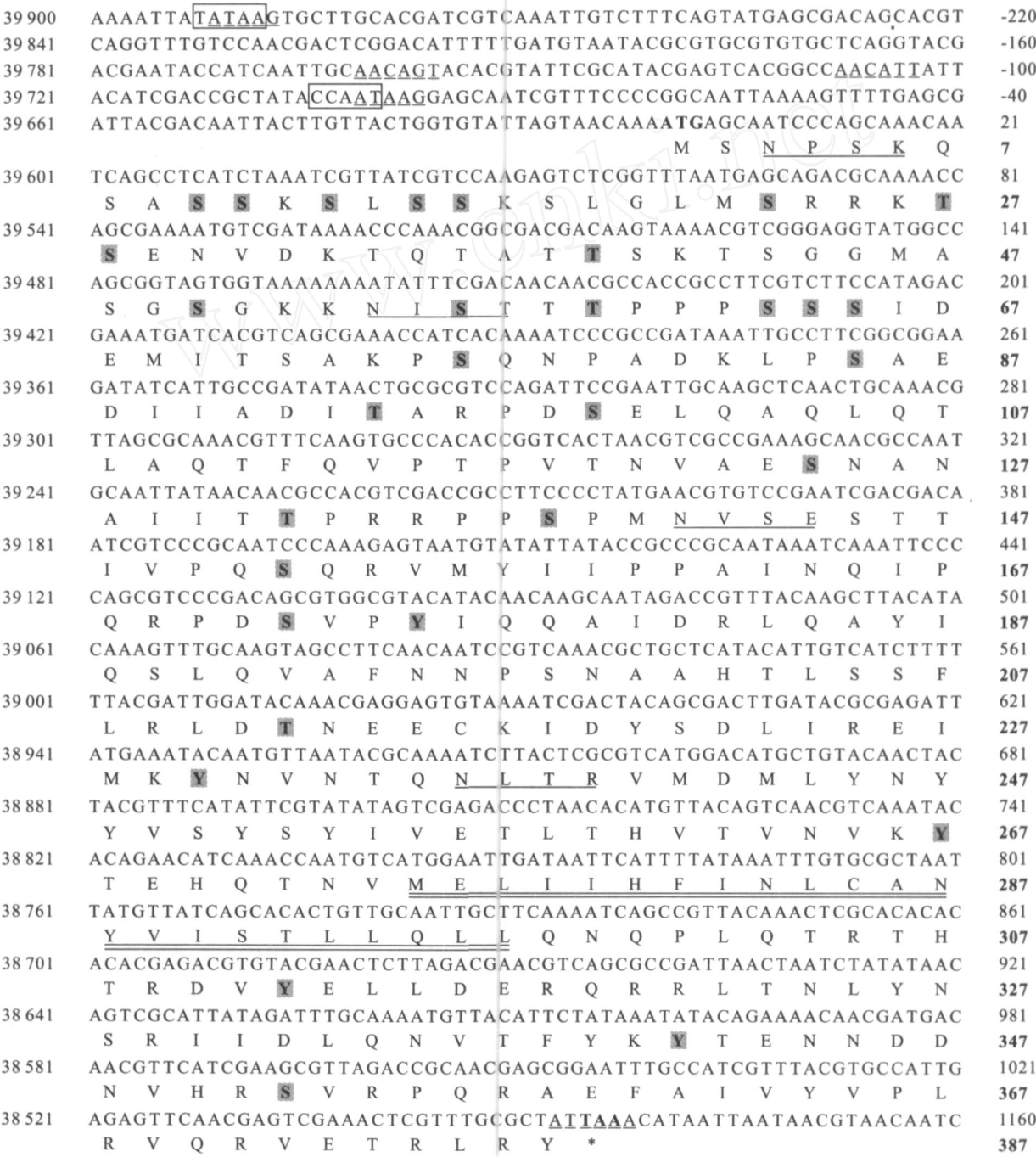


图 1 HearNPV ORF44 基因序列分析结果

黑体为起始密码子与终止密码子;虚下划线为杆状病毒早期、晚期基因转录启动子, Poly A 加尾信号;方框为 CAA 框和 TATA 框; 双下划线为跨膜螺旋区域;单下划线为糖基化位点;黑体加阴影部分为磷酸化位点

Fig.1 Sequence analysis of HearNPV ORF44

The start and stop codons are denoted in bold faces;location of baculovirus consensus early and late start sites,as well as the polyadenylation signal sequence ,are marked with broken lines;CAA and TATA boxes are framed;putative transmembrane helices region is double-underlined;putative glycosylation sites are underlined;putative phosphorylation sites are both bolded and shaded.

序列分析发现,ORF44 具有双相启动子特征序列,同时兼有杆状病毒早期和晚期基因转录启动子特征序列:起始密码子上游 103 和 136 nt 处有两个典型的杆状病毒早期基因转录启动子特征序列 AACATT 和 AACAGT;起始密码子上游 87 和 268 nt 处各有一个典型的杆状病毒晚期基因转录启动子特征序列 ATAA G。起始密码子上游 80 nt 处有一个 CAAT 框,269 nt 处有一个 TATA 框,分别与上述两个杆状病毒晚期基因转录启动子特征序列 ATAA G 重合;终止密码子 TAA 处重复一个 Poly A 加尾信号(图 1)。

用 ScanProsite、Motifscan、SignalP 分析推测氨基酸序列,未发现信号肽及其他保守基序;Tmpred 分析发现,Met275~Leu297 处有 1 个强跨膜螺旋区域;应用 NetNGlyc 分析发现 6 个 N-X-S/ T 糖基化序列,其中 4 个可能是糖基化位点;应用 NetPhos 2.0 Server 分析发现,ORF44 推导氨基酸序列中含有 32 个磷酸化位点:21 个丝氨酸激酶磷酸化位点、6 个苏氨酸激酶磷酸化位点以及 5 个酪氨酸激酶磷酸化位点(图 1)。

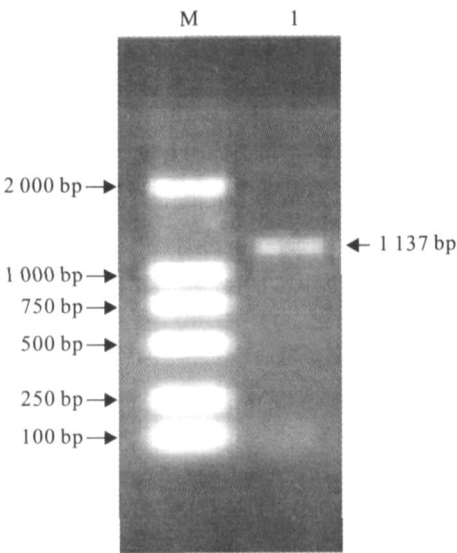


图 2 HearNPV ORF44 基因的 PCR 扩增产物
M. DNA 标准 DL2000;1. HearNPV ORF44 PCR 扩增产物

Fig. 2 PCR amplification of HearNPV ORF44

M. DNA Marker DL2000;1. PCR product of HearNPV ORF44

2.2.2 诱导表达 表达载体 pET-44 在 BL21 中诱导表达 5 h 后收集菌体,表达产物进行 12 %聚丙烯酰胺不连续凝胶电泳分析,结果如图 4 所示。与对

通过在 GenBank 和 SWISS-PROT 数据库中搜索发现,与 HearNPV ORF44 相似的基因有 10 个核型多角体病毒的 11 个 ORF (<http://www.ncbi.nlm.nih.gov/BLAST/Blast.cgi>)。将这 11 个 ORF 和 HearNPV ORF44 用 DNAMAN 进行多重序列比对发现,HearNPV ORF44 与美洲棉铃虫核型多角体病毒(*Helicoverpa zea nucleopolyhedrovirus*, HzNPV) ORF45 (NP_542668) 的同源性为 99.5 %,暗示 HearNPV 和 HzNPV 很可能是同一病毒的不同株系^[12],但与其他 10 个 ORF 的同源性仅为 16.5 %~27.5 %,表明 ORF44 是 HearNPV 的一个特有基因。

2.2 HearNPV ORF44 基因原核表达

2.2.1 重组质粒的构建 用 PCR 方法从 HearNPV 基因组中扩增 ORF44,扩增片段大小为 1 137 bp,PCR 扩增产物及 pET-44 双酶切鉴定结果分别如图 2,3 所示,结果表明片段大小与预期一致。同时,测序结果也正确无误,说明 ORF44 已正确插入到 pET-28a(+) 中。

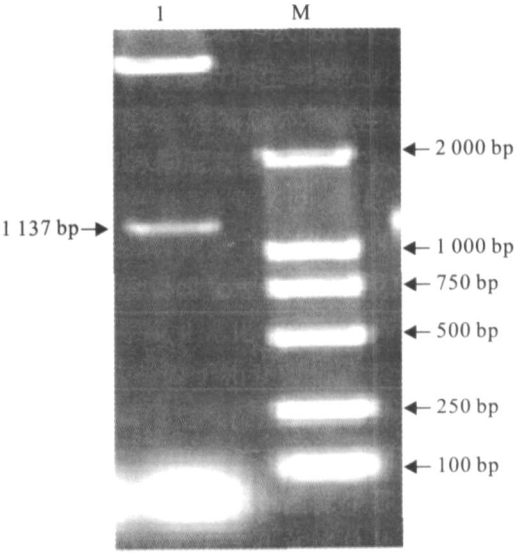


图 3 表达载体 pET-44 的双酶切鉴定结果
M. DNA 标准 DL2000;1. pET-44

Fig. 3 Restriction endonuclease digestion of expression vector pET-44

M. DNA Marker DL2000;1. pET-44

照相比,pET-44 在 44 ku 左右出现 1 条特异性高浓度条带,与预计的 ORF44 与 6 ×His 融合表达蛋白的分子量吻合。

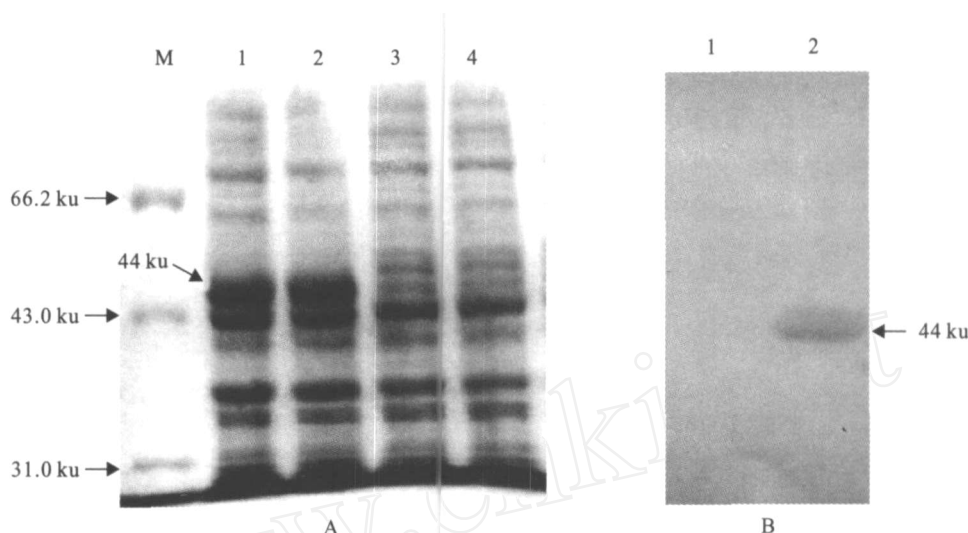


图 4 pET-44 在大肠杆菌 BL21 中的表达

A:M. 标准分子量蛋白;1,2. pET-44;3. pET-28a(+);4. BL21

B:1,2. BL21 和 pET-44 的 Western blot 分析

Fig. 4 Bacterial expression of pET-44 in E. coli (BL21)

A:M. Protein marker;1,2. pET-44 transformed BL21;3. pET-28a(+) transformed BL21;4. BL21

B:1,2. Western blot for BL21 and pET-44

3 讨 论

与 HearNPV ORF81 相似^[11], HearNPV ORF44 也具有双向启动子特征序列,不同的是,杆状病毒晚期基因转录启动子特征序列 ATAAAG 有两个拷贝,并且分别与启动子元件 TATA 框和 CAAT 框重合,但有关该元件组合对该基因转录的调控作用还需进一步深入研究。在 GenBank 和 SWISS-PROT 数据库中搜索发现,所有杆状病毒中无与 HearNPV ORF44 同源性高于 30% 的基因。在推测的 ORF44 蛋白氨基酸序列中,包含有 1 个强跨膜螺旋区域,4 个潜在的糖基化位点以及 32 个磷酸化位点,表明 ORF44 编码的蛋白可能是 HearNPV 特有的一个结构蛋白。

本研究在 BL21 中融合表达了 HearNPV ORF44,为进一步对 ORF44 编码的蛋白进行结构和功能分析等深入研究,提供了基础数据。

[参考文献]

- [1] 张光裕,刘敦化,王 锋,等. 使用棉铃虫核型多角体病毒杀虫剂防治辣椒烟青虫[J]. 中国蔬菜,1992(8):19-20.
- [2] 张传溪,武家才. 棉铃虫核型多角体病毒基因组结构及 p10 基因[J]. 生物化学与生物物理学报,2001,33(2):179-184.
- [3] Zhang C X, Ma X C, Guo Z J. Comparison of the complete genome sequences between C1 and G4 isolates of the *Helicoverpa armigera* single nucleocapsid nucleopolyhedrovirus [J]. Virology, 2005, 333(1):190-199.
- [4] 王 根,张传溪,贡成良,等. 棉铃虫核型多角体病毒多角体蛋白基因的克隆和核苷酸序列分析[J]. 病毒学报,1997,13(1):82-87.
- [5] Guo Z J, An S H, Wang D, et al. Characterization of Ha29, a specific gene for *Helicoverpa armigera* single-nucleocapsid nucleopolyhedrovirus[J]. Journal of Biochemistry and Molecular Biology, 2005, 38(3):354-359.
- [6] Wang D, An S H, Guo Z J, et al. Characterization of *Helicoverpa armigera* nucleopolyhedrovirus orf33 that encodes a novel budded virion derived protein, BV-e31 [J]. Archives of Virology, 2005, 150(8):1505-1515.
- [7] An S H, Wang D, Yang Z N, et al. Characterization of a late expression gene, Open reading frame 128 of *Helicoverpa armigera* single nucleopolyhedrovirus[J]. Archives of Virology, 2005, 150(12):2453-2466.
- [8] An S H, Shang J Y, Liu Y H, et al. Characterization of a unique gene ORF135 from *Helicoverpa armigera* single nucleocapsid nucleopolyhedrovirus[J]. Virus Genes, 2006, 32(1):21-26.
- [9] Wang D, Zhang C X. HearSNPV orf83 encodes a late, non-structural protein with an active chitin binding domain[J]. Virus Research, 2006, 117(2):237-243.
- [10] 萨姆布鲁克 J, 弗里奇 E F, 曼尼阿蒂斯 T. 分子克隆实验指南[M]. 2 版. 金冬雁, 黎孟枫, 侯云德, 等译. 北京:科学出版社, 2002.
- [11] 王 敦. 棉铃虫核型多角体病毒 (HearSNPV) ORF33、ORF80、ORF81 和 ORF83 基因分析[D]. 杭州:浙江大学, 2005.
- [12] Chen X, Zhang W J, Wong J, et al. Comparative analysis of the complete genome sequences of *Helicoverpa zea* and *Helicoverpa armigera* single-nucleocapsid nucleopolyhedrovirus [J]. Journal of General Virology, 2002, 83(3):673-684.