

## 棉铃虫核型多角体病毒 Hind III-K 片段的 核苷酸序列及其分析\*

张传溪<sup>1</sup> 孙建新<sup>2</sup> 姜育蕾<sup>2</sup> 胡萃<sup>1</sup> 吴祥甫<sup>2</sup>

<sup>1</sup>(浙江农业大学应用昆虫学研究所, 310029 杭州)

<sup>2</sup>(中国科学院上海生物化学研究所, 200031 上海)

3476.14

**摘要** 测定了棉铃虫(*Helicoverpa armigera*)核型多角体病毒(HaSNPV)基因组 DNA 的 Hind III-K 片段核苷酸序列。该片段全长 3 255 bp, 含可编码大于 40 个氨基酸残基的多肽的开放阅读框(ORF)15 个, 包括多角体蛋白(ph)基因编码区 3'端 489 bp 和蛋白激酶 HavPK 基因编码区 801 bp。在 ph 和 HavPK 两基因之间鉴定出一个可编码 412 个氨基酸残基的 ORF1 236, 转录方向与 ph 和 HavPK 基因相反。同源分析表明, ORF1 236 与谷实夜蛾(*Helicoverpa zea*)核型多角体病毒(HzSNPV)的 ORF8 推导的蛋白氨基酸序列有 95.9% 同源性, 与苜蓿丫纹夜蛾(*Autographa californica*)核型多角体病毒(AcMNPV)ORF1 629 只有 24.8% 同源性, 但三者均含有二组由多个脯氨酸残基串联而成的特征基序。

**关键词** 棉铃虫, 核型多角体病毒, 3.2 kb HindIII-K 片段, 核苷酸序列, ORF 1 236

棉铃虫[*Helicoverpa (Heliothis) armigera*]是棉花、玉米、烟草、番茄、小麦等经济作物的大害虫。化学农药的大量使用, 使生态平衡遭受破坏, 棉铃虫抗药性剧增。近年棉铃虫在我国广大棉区猖獗成灾, 造成减产甚至绝收。棉铃虫核多角体病毒(HaNPV)是我国第一个商品杀虫剂, 具使用安全、害虫不产生抗性等优点, 但亦存在杀虫致死速度较慢等缺点。因此, 通过遗传工程构建毒力强、即效性高的重组 HaNPV 具有重要的意义。对 HaNPV 进行分子生物学研究, 是今后进行其重组研究的必不可少的基础。棉铃虫单粒包埋型核多角体病毒(HaSNPV)基因组为环状双链 DNA, 整个基因组大于 120 kb, Hind III-K 片段, 大小约为 3.2 kb<sup>[1]</sup>。本文报道 HaSNPV Hind III-K 片段的核苷酸序列并对 ORF1 236 进行分析。

### 材料与方 法

#### 1 材料

HaSNPV 湖北株由中国科学院武汉病毒研究所赠送, 棉铃虫及人工饲料由中国科学院上海昆虫研究所提供, 限制性内切酶、连接酶及低熔点琼脂糖胶均为 GIBCO-BRL 公司产品, <sup>35</sup>S-dATP 为 Amersham 公司产品, 载体为 pUC19、pSK, 宿主菌为 TG1。

#### 2 方法

病毒 DNA 抽提方法见前文报道<sup>[1]</sup>, 琼脂糖分离片段、连接转化、质粒制备与纯化、限制性内切酶反应、

收稿日期: 1997-07-07, 修回日期: 1997-08-19

\* 国家自然科学基金资助项目

\*\* 本文报道的核苷酸序列数据已为 GenBank 接受, 编号为 U95055。

phagemid 单链 DNA 模板的获得均按文献[2]方法进行,单链和双链序列测定按 U. S. B. 公司的 T7 测序手册进行。ORF 分析应用 GCG package (Wisconsin Univ.) 程序,核苷酸序列和氨基酸序列进行比较应用 MALIGN 程序。

## 结果与讨论

### 1 Hind III-K 片段的核苷酸序列和 ORF 分析

在 HaSNPV 基因组 DNA 酶切图谱分析和 Ph 基因定位基础上<sup>[1]</sup>,将基因组 DNA Hind III 酶解的 3.2 kb K 片段克隆于 pUC19 质粒载体中,并用 pUC19 和 pSK 载体进行了一系列亚克隆,制备双链和单链模板,采用双脱氧链终止法测定了该片段的核苷酸全序列。图 1 为 Hind III-K 片段的限制性酶切图谱、测序策略和 ORF 分析。图 2 为该片段的核苷酸全序列,共有 3 255 bp,按杆状病毒基因组 DNA 序列惯例,以 Ph 基因编码链为正义链排列。由计算机分析表明,Hind III-K 片段中可编码至少 40 个以上氨基酸残基多肽的开放阅读框(ORF)共有 15 个。该片段包含了 Ph 基因 3 端的 489 nt 和 HavPK 基因全部编码序列 801 nt<sup>[3]</sup>,两基因转录方向相同。在 Ph 和 HavPK 基因之间,存在一个可编码 412 氨基酸多肽的 ORF 1 236,其转录方向与 Ph 和 HavPK 基因相反,即存在于 Ph 和 HavPK 的反义链上。由图 1 还可看出,在 HavPK 基因下游,还存在二个较大的 ORF,一个为 ORF228,与 HavPK 方向相同、可编码 76 氨基酸残基的多肽;另一个 ORF 延伸至片段以外,称为 ORF > 458,可编码大于 152 个氨基酸残基的多肽。这二个 ORF 能否被转录有待于进一步研究。

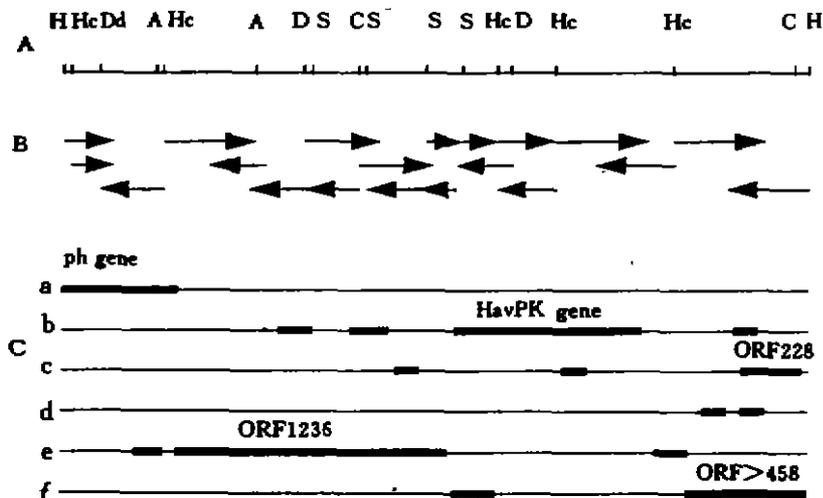


图 1 HaSNPV(湖北株)基因组 DNA Hind III-K 片段的限制性酶切图谱和 ORF 分析

Fig 1 Restriction map, sequencing strategy and ORF analysis of the HaSNPV (Hubei isolate) HindIII-K fragment

(A) Restriction map; (B) Sequencing strategy;

(C) ORF analysis of the fragment in all six reading frames: a, b and c from the polyhedrin gene coding strand of DNA, d, e and f from the opposite strand. The black boxes indicate ORFs that are at least 40 amino acids in length. The polyhedrin (Ph) gene, Protein kinase (HavPK) gene, ORF1 236, ORF228 and ORF > 458 are indicated above the boxes.

Abbreviation of Restriction enzymes: A: AvaII, D: DraI, Dd: DdeI, C: ClaI, H: HindIII, Hc: HincII, S: SalI

```

HindIII HincII
AAGCTTGTAGTTAACTGGAGCGGTCCGCAATTCTTCGCGAACTGGACTCGTTTCATGGAAGACAGTTTCC 75
+256 Polyhedrin gene-----> ClaI
ATTGTAACCGACCAAGAAATTATGGACGTGTTCTGTCTGTTAATATGCGRCAACCAACCGAACCCTTGTAT 150
CGATTCTTACCGCAACCGCTCTGCGTGTGATCCCGACTATATTCCTCACGAAGTCATTCGTATGTAGACCT 225
TCCTATGTAGGAAGTAACACCGAGTACAGAAATAGTTTAGCCAAAAATACGGCGGTGTCCCGTATGAATTTG 300
CACTCGGAATACACTAAATTCCTTTGAAGATTTTATTACCAACGTAATTTGGGAGAACTTTACAAACCAATTGTT 375
TACGTAGGCACTGATTCTGCGCAAGAGAGAAATACTCCTAGAAGTTTCTTTGATATTTAGATCAAGAATTT 450
AvaII HincII
GCACCTGACGCTCCGCTATACAGTGGTCTGCATATACTAACTTGGCATTGAGTGCATCGTCAATTTGTACTC 525
-----> polyhedrin gene stop codon***>
<*** V Q S E T S M T L K N S
ATAATTTTCTAAATTCGATCTCAATCTTGTAAATTTTIGATTGGTGGTGGTGGTGCCTAATGCCGACCCACA 600
M I K O L N S R L E Q L K Q N T P K T G L A S G V
TTAGCTAACGCTTTATCGTATTTGTTTTGAATGTCAAATCTCCACCGCCATAATGAAACCGCTAAATTTATCG 675
N A L A K D Y Q K K F T L D E V A M I F G G L N D
GACAAATGAAGTTCGACATTCACGGATTTGTCCAAAGGATATCATACTGTTGGTCAAGTGGCCGCAATA 750
S L Q L E V N V S K O L P N D Y T Q Q I L N A A I
AATATTTGTAGTTTCGCAAAATTTGTTGTTGTCGTAATTCGAAAGCGGTGATTTAATTTGATTTTATTGAT 825
F I Q L K A S I Q Q T Q A N S L R Q N L Q N K I S
ATTAATGTGCTTGTGCTTCAGTAGACAAGAAATATTTTATCCATGAACCTCCCAATGTTATATTGTACAAA 900
I L T O Q A E T S L S Y N K I W S S E L T I N Y L
AvaII
GAACGTACATATGTTTTAAATTCGCTGCTGGCTCGTCTGCTGTTGTTGCTGCTGGTCCCTCCCGTTTCCGATTCT 975
S R V Y Q K L E S S A R Q Q Q E D D T R G N E S E
HincII
GACGAAACTACAGGACTCGGTTGACGGGTATACGTCGTTGTAATCTTTTCAGTAGGACTGGCGGGCGGCGTA 1050
S S V V P S P Q R A I R R Q L I K A T P S A A A T
ACGGTATTTACTATCGAGCGGCATCGGGTTTGGATCTTTTAAATTTGATTCCTTTCTGTTATTTGTTCCATC 1125
V T N V I S R G D P K S V K K L K I G K Q I Q E M
DraI SalI
AATTCCGTCAGTGGGTCTTTTAAACTTGCAGTCCGAGTGTGATAATCCGATCTTTACTGGATGTTACT 1200
L E T R P D K L V Q R T S T T Y D R D K S S P I V
ATATCTCTATTAATGTAATGACGTTGGCGGAGGCGCGCGGAGGATCGTGAAGATAAGTTTGT 1275
I D E I L P L S P P P P P P P P P I T S S L N T
TGAAGCGCGCGGTATTTGGTGGTGGTATTTGGTGGTGGTATTTGGTGGCGCATATGTTTGGCGCGAGGAAT 1350
Q L P P P I P P P I P P P I P P P M H T Q P S S S
ClaI SalI
TCGAATTCGATAATTATAGTTGGCGAAATGTTTTTGCATTTATCCGATGTCGACACAGTTGTCGGTTAGGT 1425
E S O I I I T P S I T K Q M K D S T S V T T P K P
ATTGTTGTTTTAGTACTGTTGGTACTGACATGCTGTGCAATGTTGGTATAAATAATTGATCATTACCAATG 1500
I T T K P V T P V S M T Q S L T P L I I S R D G I
TCTATTAGTACGCTGTTGTTGATATTTCTTGGGCAATTTCAATAACTGAATCAATCGTATACGCTAATTTGT 1575
D I L V D N N Y I E Q A L K L L Q I C Y V D L Q
ATCCGATCAGAATTGGACTGAGCGCAGCGGTGACCGTACGTTTCAAACTGCGGGCGCGAGATCAAGCGCAAT 1650
I R O S N S Q A V A S V T R K L S H P A S I M R L
SalI
AGAAAGTCGACATTTATGATGTTGTGTAGTTTTTTTTCAGCCAAATATGTTGAACACTTTGCAAGTTGAACATT 1725
L F D V N N I N T Y N K E A L Y Q Q V S Q L Q V M <
ATTGCGAATCGCAGGACGACCGGTTTCGTTAAGGAAATAAACCAATTTTTCGCGCAATAAATAACAAACAA 1800
SalI >-----> HavPK gene start
TGTGCGTTTGGTCCGCGCAAGTTTGGCAAAATGTGTGTTATCAAACACGAGCCCAACCGGCAAACTGTTCCGAAA 1875
AAAGAGTCCCAATTAATATGTCACCGGATCGAACCTATGGTGCATCAACTAATGAAGGACAAACCGATTTT 1950
HincII
CATCAAATATATTAATCTGTTGACGACGTTAAATCTCAAATACTCATATTAGATTACGTTGCTGGAGCGGATTT 2025
DraI
GTTTGATTTTTTAAAAAACCAAAAAAGTATCTGAAGCGGAAACCGTTCATAGTGGTCAATTAACCGAAGC 2100
ACTGAAGCGGCTTCACTCTTACAAAATTATACATAACGATCTCAAACCTCGAAAACGTCCTATACGTACGTATAA 2175
HincII
ACAAATTAATTTGTGTGATATGGACTGTGTAATTTTCAACACGAGTTCGTTGTCGAGACGGCCAAAAGGAGTA 2250
CATGCTCCGGAGAGCTCAACGACAAAACCTACGATGTTCAACGTCGATTTGGTGGCTTTGGGCACTTTGACGTA 2325
TGAACCTTTAATTTGGACATCATCCCTACAACATAGCABCACACGATGAAGATTTCGATTTGGATGACTACA 2400
ACAGAGACAAACAAAAAACTTTACAATACAATTTTTAAGTAGTACGCTCAAAAATTTTGGAAACCAATGTT 2475
AATGTATAACATTAATTACAGGTTGTGTACATACGAGACTGTAATAAACACAGTTTTTTTACAATAATATATT 2550
HavPK gene translational stop codon>***
TAATAAAAAAATAATGTTTTTTTATACCAATCAACTAAAGTATAAATAATACAAAAGTTTATTACAA 2625
TCAATTCAACTAATATATGATATACAAAAGTTACATTAATATATCTGCTTTGGGAGCAGCTGAAGTGGC 2700
HincII
TTGACGTTTAGCTGGTGGTTTTTCAGTACGAGAACTGGTACTCTAACCATACGAAAAGTAGCTATCTGAGGTTTC 2775
ATGTTATCTCCCAATGCACTATTTCACCTCATCTGTCATCTGTCATTCAGCAACTAGCAGGGCTTAAAGGTA 2850
***<ORF>458_stop codon<-----
AATTCAAACATTCAACATCAGACATATCGACAGGTTCTTGTTTGGGAACACATCTTCATGATACTCATTAAAT 2925
AATCAGCATTTTTCACATTAGTATTAATATATCCCAACCAATCTTTTTTATGCAATGTCAAATGGCGCAG 3000
> ORF228 start-->
TGTCAATTAATATCAGTAGTGTAGCATCCTTTGATGTTTTTCTGTTTTTAAACAATAATGCTCGAAATATT 3075
ClaI
TGCCATTTTGTCTACATTGCTACTTTTACCTAATCTTTATCGATACTATCAAGTTCTTCAGTACTCATTGCAA 3150
CTGGTGTACTGTGTTGATGATGTTCTTTTTCAAGCAATTTGCGCACTTCATTTCAATTTGACTTATTTCTG 3225
HindIII
TCAATTTGACACAAGTACTTCTGAGCTT 3255

```

图2 HaSNPV 3 255 bp HindIII-K 片段的核苷酸序列和由 ORF1 236 推导的氨基酸序列

Fig 2 Nucleotide sequence of HaSNPV 3 255bp HindIII-K fragment and predicted amino acid sequence of the ORF1 236

The nucleotide sequence is given in the genomic DNA orientation direction as the transcriptional direction of the polyhedrin gene.

ORF1 236 is on the opposite strand of the DNA. Relevant restriction enzyme sites are underlined and indicated above the line. Numbers at the ends of the line refer to nucleotide sequence. The translational start codon and stop codon are in bold for polyhedrin gene, *HavPK* gene or ORF228, and in italics for ORF1 236 or ORF>458.

## 2 ORF 1 236 核苷酸序列特征

从图 2 可见, ORF 1 236 起始密码子 ATG 位于 1 724 nt 处; 终止密码子 TAA 位于 488 nt 处, 并与 Ph 基因终止密码子 TAA 发生相遇且部分重叠。为了便于进行序列分析, 将 ORF 1 236 的启动子区和部分编码区编码链序列列于图 3。由图 3 可见, 在 ATG 上游、-398 nt 处隐含杆状病毒晚期基因典型的启动子特征序列, 即 Rohrmann box 核心序列 GTAAG, 而且启动子区富含 AT, 约占 70%; 同时 5' 端侧翼区含有 3 个串联重复的 CAGT, 分别出现在 ATG 上游-143、-380 和-479 nt 处。CAGT 为杆状病毒早期基因启动子基序特征; 此外还具有真核生物启动子特征序列 TATA box(位于-238 和-407 nt)和两个 CAAT(-489 nt, -5 nt)特征序列, 而且含有真核生物 mRNA 高水平表达的调控信号 Kozak 序列 AUAAUGG, -3 位仍具有保守碱基 A。该 ORF 近处没有出现明显的 Poly(A)终止信号序列, 可能在较远端。

```

HincII
GTTGCAATTTTACACAGTCCATAATCACACAAATAAATTTGTTTATGACGTACGTATAG 60
CACGTTTTTCGAGTTTGAGATCGTTATGTATAATTTGTAAGAGTGAAGCGCGTTCAGTGC 120
TTCGGTTAATTGACCCACTATTGAACGTGTTTCCGCTTCAGATACTTTTTTGTGTTTTTT 180
TAAAAAATCAAACAAATCGCCTCCAGCAACGTAATCTAATATGAGTATTTGAGATTTTAA 240
HincII
CGTCGTCAACGAGTAATTATAATTTGATGAAATATCGGTTGTCCTTCATTAGTTGATGCAC 300
CATAGGTTTCGATCTCGGTCACATATTTAATTGCGACACTCTTTTTTACGAACAGTTTGCC 360
Sall
CGTGGGCTCGTGTTTGATAACACACATTTTGCCAAACTTGCCGTCGACCAAACGCACATT 420
GTTTTGTATTTTTATTTTCGGCGAAAAATTGGTTTATTTCCGGAACGAAACGGTTCGTCAT 480

TGCGATTTCGCAATATATGGTTCAACTGCAAAGTGTTCAACAATATTTGGCTGAAAAA... 540
ORF1236----> M V Q L Q S V Q Q Y L A E K ...

```

图 3 HaSNPV ORF1 236 5' 端非编码区的核苷酸序列

Fig 3 Nucleotide sequence of the 5' flanking region of HaSNPV ORF1 236

Eukaryotic consensus CAT and TATA boxes are written with italic bold and bold letters, respectively;

Motifs for early (CAGT) and late (GTAAG) NPV gene promoters are in italics and underlined; The enzyme sites are single underlined; The arrows denote the translational directions of ORF1 236.

## 3 同源性分析

ORF1 236 在 HaSNPV 基因组中的线性位置与苜蓿丫纹夜蛾 (*Autographa californica*) 核型多角体病毒 (AcMNPV) 的 ORF1 629 和谷实夜蛾 (*Helicoverpa zea*) 核型多角体病毒 (HzSNPV) 的 ORF8 在各自基因组中的线性位置完全相同, 都位于多角体蛋白和蛋白激酶两基因之间, 并与两基因紧邻, 且编码方向都与两基因相反。将所测的 ORF1 236 核苷酸序列和推导的氨基酸序列与 AcMNPV ORF 1 629<sup>[4]</sup> 和 HzSNPV ORF 8<sup>[5]</sup> 进行比较, 结果列于图 4。HaSNPV 和 HzSNPV 的编码区核苷酸、氨基酸序列非常相似, 二者之间核苷酸同源性达 96.4%, 氨基酸同源性达 95.9%。在所编码的 412 个氨基酸中只有 17 个不同, 其中在图 4 的第 212 处 HaSNPV ORF 1 236 比 HzSNPV ORF 多一个亮氨酸(L), 在第 492 处少一个赖氨酸(K)。但 HaSNPV ORF 1 236 与 AMNPV

的 ORF 1 629 同源性较低,启动子区和编码区的核苷酸序列同源性分别为 49.5% 和 32.8%。AcMNPV ORF 1 629 编码有 543 个氨基酸残基,比 HaSNPV ORF 1 236 和 H<sub>2</sub>SNPV ORF8 编码的 412 个氨基酸残基长 131 个,两者氨基酸序列只有 24.8% 同源性。

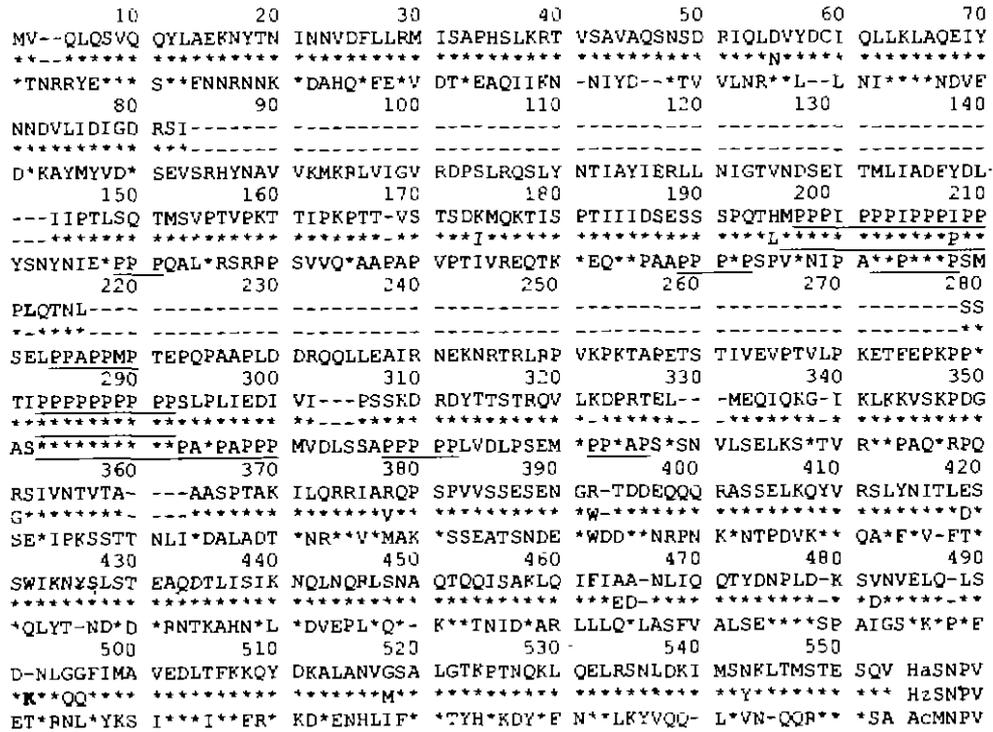


图 4 HaSNPV ORF1 236 推导的氨基酸序列(上行)和 H<sub>2</sub>SNPV ORF 8(中行)及 AcMNPV ORF 1 629(下行)的氨基酸序列联配

Fig 4 Alignment of amino acid sequences for the predicted ORF 1 236 of HaSNPV (upper line), ORF 8 of H<sub>2</sub>SNPV (middle line) and ORF 1 629 of AcMNPV (lowest line)

The insertion or gap residues between ORF 1 236 and ORF 8 are in bold.

但是, HaSNPV ORF 1 236 和 H<sub>2</sub>SNPV ORF 8 及 AcMNPV ORF 1 629 有一个共同特征,即均由多个脯氨酸(P)残基串联在一起的基序,特别是存在两组脯氨酸串联,中间由若干个其它氨基酸链相隔形成的明显保守区域(图 4),脯氨酸占整个编码蛋白的 9.5~15%,在甘蓝夜蛾(*Mamestra brassicae*)核型多角体病毒(MbMNPV)<sup>[6]</sup>和茶尺蠖(*Ectropis obliqua*)核型多角体病毒(EoSNPV)Ph 基因下游均发现有串联的脯氨酸残基(张传溪等,未发表资料)。因此,这一保守区可能对其功能有重要作用。

已知 AcMNPV ORF 1 629 在病毒感染细胞后 12 h 开始转录,到 48 h p.i 达到高峰<sup>[7]</sup>。ORF 1 629 蛋白可能是 PDV 和 BV 病毒粒子囊膜或多角体膜的一个组分,也可能具有潜在的跨膜作用<sup>[8]</sup>。通过对 AcMNPV 缺失突变发现,位于多角体基因下游的 ORF 1 629 蛋白可能对于病毒复制是必需的<sup>[4]</sup>。对 HaSNPV 的 ORF1 236 有待于进行转录和功能分析。

## 参 考 文 献

- 1 王根,张传溪,金伟等.棉铃虫核型多角体病毒限制性酶切图谱和多角体蛋白基因定位.环境与应用生物学报,1996,2(4):387-394
- 2 Sambrook J, Fritsch E F, Maniatis T *et al.* Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory Press, NY; Cold Spring Harbor, 1989
- 3 张传溪,王根,胡萃等. HaSNPV 蛋白激酶基因的核苷酸序列分析.生物物理与生物化学学报,1997,29(3):322-326
- 4 Pham, D P Q, Sivasubramanian N. Sequence and *in vitro* translational analysis of a 1 629-nucleotide ORF in *Autographa californica* nuclear polyhedrosis virus strain E2. Gene, 1992, 122:345-348
- 5 Cowan P, Bulach D, Goodge K *et al.* Nucleotide sequence of the polyhedrin gene region of *Helicoverpa zea* single nucleocapsid nuclear polyhedrosis virus; placement of the virus in lepidopteran nuclear polyhedrosis virus group II. J Gen Virol, 1994, 75(11):3211-3218
- 6 Cameron I R, Possee R D. Conservation of polyhedrin gene promoter function between *Autographa californica* and *Mamestra brassicae* nuclear polyhedrosis viruses. Virus Research, 1989, 12:183-199
- 7 Ooi B G, Miller L K. Transcription of the baculovirus polyhedrin gene reduces the levels of an antisense transcript initiated downstream. J Virol, 1990, 64(6):3126-3129
- 8 Brunagel S C, Summers M D. *Autographa californica* nuclear polyhedrosis virus, PDV, and ECV viral envelopes and nucleocapsids: Structural proteins, antigens, lipid and fatty acid profiles. Virology, 1994, 202:315-328

## Nucleotide Sequence of the *Helicoverpa armigera* Nuclear Polyhedrosis Virus HindIII-K Fragment and Its Analysis\*

Zhang Chuanxi<sup>1</sup> Sun Jiangxin<sup>2</sup> Jiang Yulei<sup>2</sup> Hu Cui<sup>1</sup> Wu Xiangfu<sup>2</sup>

<sup>1</sup>(Institute of Applied Entomology, Zhejiang Agricultural University, Hangzhou 310029)

<sup>2</sup>(Shanghai Institute of Biochemistry, Academia Sinica, Shanghai 200031)

**Abstract** Nucleotide sequence of the *Helicoverpa armigera* single nucleocapsid nuclear polyhedrosis virus (HaSNPV) HindIII-K fragment has been sequenced. The fragment comprises 3 255 bp, containing 15 ORFs that are at least 40 amino acids in length, including 489 bp of the 3' end of polyhedrin (ph) coding region and complete coding region (801 bp) of the protein kinase (HavPk). An open reading frame, ORF 1 236, encoding a 412 amino acid polypeptide was identified between immediately downstream ph gene and upstream HavPK gene, but the transcriptional direction was opposite to them. Comparison of the nucleotide sequences and deduced amino acid sequences demonstrated that the ORF 1 236 is closely similar to the ORF 8 of *H. zea* single nucleocapsid nuclear polyhedrosis virus (HzSNPV), with 95.9% amino acid identity, but is distantly related to the ORF 1 629 of *Autographa californica* multiple nucleocapsid nuclear polyhedrosis virus (AcMNPV), with only 24.8% amino acid identity. The three proteins were characterized by the presence of two groups of multiple tandem proline residues.

**Key words** *Helicoverpa armigera*, Nuclear polyhedrosis virus, 3 255bp HindIII-K fragment, Nucleotide sequence, ORF1 236

\* The nucleotide sequence data reported here have been deposited with the GenBank database and assigned the accession number U95055.