

棉铃虫单核衣壳核多角体病毒解螺旋酶基因的序列分析*

范凌云¹, 胡志红¹, Just. M. Vlak², 陈新文^{1**}

(1. 中国科学院武汉病毒研究所无脊椎动物病毒学联合实验室, 湖北武汉 430071; 2. Laboratory of Virology, Wageningen University, Bovenlaan 11, 6709PD Wageningen, Netherlands)

Sequence Analysis of *Helicoverpa armigera* Single-nucleocapsid Nucleopolyhedrovirus helicase GeneFAN Ling-yun¹, HU Zhi-hong¹, Just. M. Vlak², CHEN Xin-wen^{1**}

(1. Joint-lab of Invertebrate Virology, Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan 430071, China; 2. Laboratory of Virology, Wageningen University, Bovenlaan 11, 6709PD Wageningen, Netherlands)

Abstract: In order to investigate the genomic organization of the single-nucleocapsid nucleopolyhedrovirus of *Helicoverpa armigera*, the *EcoR* I-N fragment located at 54.8 ~ 59.3 kbp of the viral genome was sequenced. The fragment contained 3 762 bp *helicase* gene potentially encoding a protein with a molecular mass of 146kDa. A late transcriptional motif, ATAAG, was found at -50nt upstream of the translational start codon ATG, while two TATA box was located at -112nt and -189nt upstream of ATG. A typical poly(A) signal was found at 12nt downstream of the translational stop codon. Compared HaSNPV helicase with the helicases of *Autographa coliformica* MNPV (AcMNPV), (*Bombyx boris* NPV (BmNPV), *Lymantria dispar* MNPV (LdMNPV) *Spodoptera exigua* MNPV (SeMNPV), *Orgyia pseudotsugata* MNPV (OpMNPV), and *Xestia c-nigrum* granulovirus (XcGV), only 5 motifs (I, Ia, II, III, IV) were found conserved in baculovirus. The homology of the other two motifs (V and VI), which are also quite conserved in other helicases, were lower than 30%. The HaSNPV helicase protein had considerable amino acid sequence similarity with other baculovirus helicases (58%), and the highest (66%) with SeMNPV and the lowest with XcGV (43%). HaSNPV helicase was the first helicase reported in single-nucleocapsid nucleopolyhedrovirus.

Key words: *Helicoverpa armigera* single-nucleocapsid nucleopolyhedrovirus (HaSNPV); Helicase gene; Sequence analysis

摘要: 对棉铃虫单核衣壳核多角体病毒 (*Helicoverpa armigera* single-nucleocapsid nucleopolyhedrovirus, HaSNPV) 基因组中 *EcoR* I-N 片段进行序列分析, 获得了完整的解螺旋酶基因 (*hel*), 其开放阅读框大小为 3 762bp, 编码一个分子量为 146kD 的蛋白质。在 *hel* 起始密码子 ATG 上游 50 位有强晚期启动子转录起始信号 ATAAG, 在 -112 位和 -189 位存在两个 TATA box, 但未发现早期转录信号 CAGT。其在终止密码子下游第 12 位有一 PolyA 终止信号 AATAAA。在其它真核或原核解螺旋酶中存在的 7 个保守基元 (I、Ia、II、III、IV、V、VI), 只有 5 个 (I、Ia、II、III、IV) 在杆状病毒中保守。

收稿日期: 2000-10-09, 修回日期: 2001-02-14

* 基金项目: 中国科学院青年科学家小组基金; 国家自然科学基金 (39970031, 30025003)

作者简介: 范凌云 (1976-), 男, 湖北武汉籍, 硕士研究生, 研究方向分子生物学。

** 通讯作者. Correspondence author.

本文报道序列已收录于 GenBank, 登记号 AF266695

同源性比较发现,HaSNPV 解螺旋酶的氨基酸序列与甜菜夜蛾核多角体病毒(*Spodoptera exigue* MNPV, SeMNPV)的解螺旋酶具有最高的同源性(66%),与 *Xestia c-nigrum* 颗粒体病毒(XcGV)解螺旋酶的同源性最低(43%)。HaSNPV 解螺旋酶基因是第一个报道的单粒包埋核多角体病毒的解螺旋酶基因。

关键词:棉铃虫单核衣壳核多角体病毒(HaSNPV),解螺旋酶基因,序列分析

中图分类号:Q812

文献标识码:A

文章编号:1003-5125(2001)04-0344-06

杆状病毒是一类能感染节肢动物(主要是鳞翅目昆虫)的病毒。其病毒粒子呈杆状,基因组为双链环状 DNA,大小为 90~160kb。杆状病毒科(Baculoviridae)分为两个属:核型多角体病毒属(Nucleopolyhedrovirus,简称 NPV)与颗粒体病毒属(Granulovirus,简称 GV),其中 NPV 根据病毒粒子内包埋单个或多个核衣壳区分为单粒包埋型 NPV(SNPV)和多粒包埋型 NPV(MNPV)。杆状病毒因为具有高致病性、宿主专一性以及不易导致害虫产生抗性等优点已广泛应用于有害昆虫的防治。另外,杆状病毒作为一种高效的真核表达系统已得到广泛的应用^[1]。

通过瞬时表达实验从苜蓿银纹夜蛾核多角体病毒(*Autographa californica* MNPV, AcMNPV)基因组中鉴定出 10 个与病毒 DNA 复制有关的基因^[2],其中 DNA 聚合酶基因(*dnapol*)、晚期表达因子-1 基因(*lef-1*)、晚期表达因子-2 基因(*lef-2*)、晚期表达因子-3(*lef-3*)、解螺旋酶基因(*hel*)和极早期基因-1(*ie-1*)被确认为是杆状病毒 DNA 复制必须基因^[2]。

解螺旋酶可以破坏互补碱基对中的氢键,催化双链 DNA、双链 RNA 和 DNA-RNA 杂合体的解螺旋,对 DNA 的复制、修复、重组、转录起着重要的作用。AcMNPV 的 *hel* 编码一个 143kD 的蛋白,具有与解螺旋酶相似的功能结构域,并能与双链 DNA 结合^[3],现已有实验证明 AcMNPV P143 具有解螺旋酶的功能^[4]。此外,*hel* 基因还与杆状病毒的宿主范围有关^[5]。目前已从家蚕核多角体病毒(*Bombyx mori* NPV, BmNPV)^[6],舞毒蛾核多角体病毒(*Lymantria dispar* MNPV, LdMNPV)^[7],甜菜夜蛾核多角体病毒(*Spodoptera exigue* MNPV, SeMNPV)^[8],黄杉毒蛾核多角体病毒(*Orgyia pseudotsugata* MNPV, OpMNPV)^[9]和 *Xestia c-nigrum* 颗粒体病毒(XcGV)^[10]等 5 种杆状病毒中发现了解螺旋酶基因的存在,但尚未有关于单核衣壳核多角体病毒(SNPV)解螺旋酶基因的报道。

棉铃虫核多角体病毒(*Helicoverpa armigera* SNPV, 简称 HaSNPV)是棉铃虫的特异性病原,自 1973 年已被用于棉铃虫的防治,并已在我国开发为

商品杀虫剂,每年应用面积达 $1.0 \times 10^5 \text{hm}^2$,收到了良好的社会效益。为了构建新的高效重组病毒,我们对 HaSNPV 的分子生物学进行了广泛的研究,构建了 HaSNPV 的基因文库和物理图谱^[11],测定了该病毒的多角体蛋白基因(*polyhedrin*)^[12]、*egt*^[13]、*lef-2*^[14]等基因序列。本文报道 HaSNPV 解螺旋酶基因序列并对其结构特点进行比较分析。

1 材料与方法

1.1 多角体与病毒 DNA 的提取

HaSNPV 多角体感染健康的 4 龄初中国棉铃虫后,收集典型的病死虫虫尸,按孙修炼等的方法^[15]提纯多角体,然后经碱解,酚、氯仿抽提,透析 48h,得到纯净的病毒 DNA。

1.2 克隆及序列测定

HaSNPV 基因组 DNA 用限制性内切酶 *Eco*R I、*Hind* III 和 *Xba* I 消化并用鸟枪法克隆到质粒载体 pTZ19R 上^[11]。在对文库进行随机测序过程中,在 *Hind* III-M 片段上发现部分 *hel* 基因,随后对包含有 *Hind* III-M 中的 *Eco*R I-N 片段用特异引物采用引物步移法进行双向测序,获得了完整 *hel* 基因。

1.3 计算机分析

所测得的序列应用 UWGCG 程序(version 10.0)和 DNASTar 软件包进行分析。将测定的 HaSNPV *hel* 基因序列和推测的氨基酸序列用 BLAST 与 GenBank/EMBL 数据库进行比较。应用 Megalign 软件、ClustalX 软件和 GeneDoc 软件(Version 1.1004)进行氨基酸序列同源性比较分析。

2 结果与讨论

2.1 *hel* 在 HaSNPV 基因组上的定位

在对 HaSNPV 的 *Hind* III 基因文库的随机序列测定中,我们在 *Hind* III-M 片段发现了 HaSNPV *hel* 基因,用引物步移法以特异引物对包含有 *Hind* III-M 片段的 *Eco*R I-N 片段进行双向测序,获得 *hel* 基因的完整序列。经与 HaSNPV 物理图谱比较^[11],该基因位于基因组的 54.8Kb~59.3Kb 处,转录方向与

多角体蛋白基因方向相反(图1)。在 *hel* 基因下游还发现一个不完整的开发阅读框(ORF),其转录方向与多角体蛋白基因相反(图1)。该基因与 SeMN-

PV ORF68 的 N 端具有同源性^[8]。到目前为止该基因仅在 SeMNPV 有报道,其功能不详。

2.2 HaSNPV *hel* 基因结构

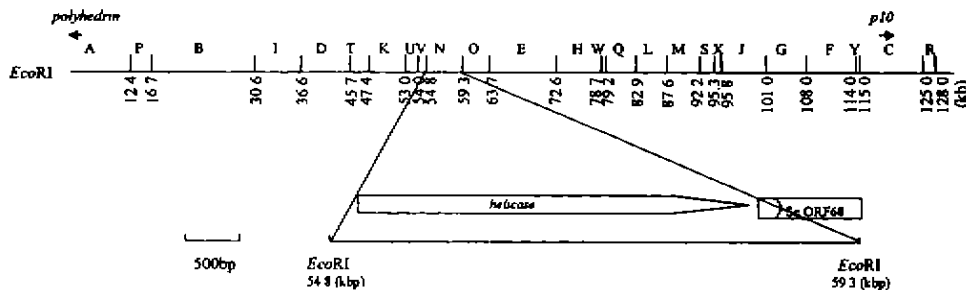


图1 HaSNPV 基因组的 *EcoRI* 酶切图谱及 *EcoRI*-N 片段的结构

Fig.1 *EcoRI* physical map and detailed gene organization of the *EcoRI* 1-N fragment.

The arrows represent the ORFs and their directions of transcription.

HaSNPV 解螺旋酶基因大小为 3 762bp, 编码一个 1 253 个氨基酸的蛋白质, 预计分子量为 146kD。在 *hel* 起始密码子 ATG 上游 50 位有强晚期启动子转录起始信号 ATAAG, 在 -112 位和 -189 位存在两个 TATA box, 但未发现早期转录信号 CAGT。除了在 LdMNPV 的 *hel* 起始密码子 ATG 上游 125 有一个早期转录信号 CAGT 外, 其它的 5 种杆状病毒也不具有早期转录信号 CAGT。在终止密码子下游第 12 位有一 PolyA 终止信号 AATAAA。HaSNPV *hel* 的转录特性有待进一步分析。

2.3 HaSNPV 解螺旋酶氨基酸序列比较分析

HaSNPV 解螺旋酶与 AcMNPV^[16], BmNPV^[6], OpMNPV^[9], SeMNPV^[8], LdMNPV^[7] 及 XcGV^[10] 等杆状病毒的解螺旋酶进行了同源性比较发现, 杆状病毒解螺旋酶具有较高的保守性, 平均同源性高达 58%, 其中与 SeMNPV 的同源性最高, 为 66%, 而与 XcGV 的同源性最低, 为 43% (表 1)。

表 1 7 种杆状病毒解螺旋酶氨基酸同源性 (%) 比较

Table 1 Amino acid sequence identity (%) of seven baculovirus helicases

	BmNPV	OpMNPV	SeMNPV	LdMNPV	XcGV	HaSNPV
AcMNPV	98	76	59	59	41	60
BmNPV		76	59	59	41	60
OpMNPV			56	56	39	55
SeMNPV				67	41	66
LdMNPV					42	64
XcGV						43

在已报道的原核和真核解螺旋酶中存在七个高度保守基元序列 (motif I, motif Ia, motif II, motif III, motif IV, motif V 和 motif VI), 而且这些基元序列 (N-末端-I-Ia-II-III-IV-C-末端) 的空间顺序也是十分保守的^[17,18]。但比较分析发现在 HaSNPV 解螺旋酶的 C 端只有五个基元序列 (motif I, motif Ia, motif II, motif

III 和 motif IV) 高度保守。Motif I (aa960 ~ 973) 含有 GTP 和 ATP 结合蛋白 A-位点^[19] 或 P-loop^[20] 的特征序列 (GxxxxGKS), 该区域在杆状病毒解螺旋酶中高度保守, 同源性高达 78%。此基元序列形成一个环, 结合着磷酸基团, 另一处通过 Mg²⁺ 结合一个天冬氨酸, Motif Ia 位于第 1 004 ~ 1 021 氨基酸, 已有的研究表明该区域是解螺旋酶与 DNA 结合的重要活性部位^[13]。HaSNPV 解螺旋酶在 motif II (1 057 ~ 1 067aa) 中并未发现典型的 NTP 结合蛋白 B-位点的特征序列 DD^[19], 而相对应的序列为 SD, 这与 SeMNPV 相同^[21]。这种变化可能引起在 P-loop、镁离子、B-位点中的天冬氨酸之间相互作用的改变^[17-21], Motif III (1 084-1 095aa) 与病毒 DNA 聚合酶的一个保守区域相似, 但 Motif III 的功能不详。Motif IV (1 121-1 129aa) 的功能还未确定。

相对于其它的解螺旋酶基因序列比较情况杆状病毒包括 HaSNPV 的解螺旋酶基因的 motif 和 Motif VI (1 121-1 129aa) 的功能还未确定。在对单纯疱疹病毒-1 型的解螺旋酶 UL-5 诱变实验中, 发现在解螺旋酶中所有的基元序列对其在体内的活性都是必需的。突变实验发现 AcMNPV 的解螺旋酶 P143 的 motifs I, Ia, II 和 III 的改变才影响了 P143 拯救温度敏感复制缺陷型突变体的能力, 而 motifs IV, V 和 VI 单一保守区的突变对 P143 的拯救能力无明显影响^[22], 且杆状病毒解螺旋酶的 motifs IV, V 和 VI 具有极低的或不具有保守性 (图 2), 说明 motifs IV, V 和 VI 对杆状病毒解螺旋酶的功能是非必需的。例如, 发生在大肠杆菌 Rep 蛋白的 motif IV 中的一个突变, 并不消除其在体外解螺旋酶的活性^[23]。

图 2 杆状病毒解螺旋酶氨基酸序列同源比较

Fig. 2 Alignment of the deduced amino acid sequences of seven helicase genes.

The seven conserved helicase motifs are underlined and denoted I through V; nuclear localization signals are indicated by *NLS*; amino acids involved in host range expansion are indicated by asterisks superscribed by 1 host range factors' and the regions in the seven baculovirus helicases that display a higher degree of identity are undlined and denoted A1 through A5.

比较分析发现杆状病毒的解螺旋酶还具有其它高度保守分区域: A1(aa160~174), A2(aa550~557), A3(aa888~900), A4(aa920~933), A5(aa1026~1038), 其中 A2、A3 和 A4 区域的同源性均大于 80% (图 2)。这些区域可能是杆状病毒的解螺旋酶所特有的^[21]。

序列分析发现 HaSNPV 存在三个可能的核定位信号(KxxK/R)(在 HaSNPV 中分别位于第 623, 629 和 833 位氨基酸)[图 2]。核定位信号起到“定向”、“定位”的作用,可保证整个蛋白质通过核孔复合体的核输入。前两个假定的核定位信号(KxxK/R)在 6 个杆状病毒的 *hel* 基因中都存在。第三个位于富含天冬氨酸的区域,是 HaSNPV 所特有的,其是否具有功能尚待进一步研究。

HaSNPV 解螺旋酶还含有类似 Leu-*zip*(I¹¹⁴-V¹²¹-I¹²⁸-L¹³⁵-V¹⁴²), 在 AcMNPV, BmNPV, SeMNPV 和 OpMNPV 中也存在类似的基元序列。大多数解螺旋酶是以寡聚体形式表现出活性的。被修饰的拉链基元序列可能与解螺旋酶的寡聚化有关。这类序列可能使 *hel* 基因产物形成同源二聚体或使 *hel* 基因产物与病毒特异蛋白形成同源二聚体^[24]。但 L-*zip* 的保守性并不高, XcGV 缺乏此特征序列(图 2), 这可能是杆状病毒解螺旋酶特异性的原因之一。

AcMNPV *hel* 基因的一个 79nt 的区段被 BmNPV *hel* 基因对应的碱基序列取代后, 所获得的重组病毒既可以感染 AcMNPV 的敏感细胞系 Sf, 也能感染 BmNPV 的敏感细胞系 BmN(该细胞系为 AcMNPV 的非敏感细胞系)^[5]。比较发现该区域在已报道的杆状病毒中保守性并不高(图 2)。其中影响 AcMNPV 宿主范围的 4 个氨基酸(在 AcMNPV 中为 H⁵⁵¹, V⁵⁵⁶, S⁵⁶⁴, F⁵⁷⁷ 高度不保守^[5])。但在 HaSNPV 中相对应的氨基酸为 Q⁵⁶³, V⁵⁶⁸, S⁵⁷⁶, F⁵⁸⁹, 与 AcMNPV 只有一个氨基酸发生变化。AcMNPV 和 HaSNPV 均能感染秋粘虫和棉铃虫, 虽然交叉感染即 AcMNPV 感染棉铃虫, HaSNPV 感染秋粘虫的感染率不高。也许这四个氨基酸的相似性与两病毒的宿主特异性相关。然而 AcMNPV 具有比 HaSNPV 广得多的宿主范围, 暗示解螺旋酶的其他位点或其它基因也参与决定杆状

病毒的宿主范围。

本文首次报道了单粒包埋核多角体病毒的解螺旋酶基因, 为进一步研究杆状病毒解螺旋酶的结构与功能及 HaSNPV 的复制规律提供了科学基础。

参考文献

- [1] Miller L K. The Baculovirus[M]. New York: Plenum Press, 1997. 1-5
- [2] Kool M, Voeten J T M, Goldbach R W, et al. Functional mapping of regions involved of the *Autographa californica* nuclear polyhedrosis virus genome required for DNA replication[J]. Virology, 1994a, 198: 680-689.
- [3] Laufs S, Lu A, Arrell k, et al. *Autographa californica* nuclear polyhedrosis virus p143 gene product is a DNA binding protein[J]. Virology, 1997, 228: 98
- [4] Vivien V, Guarino A. The *Autographa californica* nuclear polyhedrosis virus p143 gene encodes a DNA helicase [J]. Journal of Virology, 2000, 74: 5273-5279
- [5] Croizier G, Croizier L, Argaud O, et al. Extension of *Autographa californica* nuclear polyhedrosis virus host range by interspecific replacement of a short DNA sequence in the p143 helicase gene[J]. Pro Nat Acad Sci USA, 1994, 91: 48-52.
- [6] Gomi S, Majima K, Maeda S. Sequence analysis of *Bmobyx mori* nucleopolyhedrovirus[J]. J Gen Virol, 1999, 80: 1323-37.
- [7] Kuzio J, Pearson M N, Harwood S H. Sequence and analysis of the genome of a baculovirus pathogenic for *lymantria dispar* [J]. Virology, 1999, 253(1): 17-34.
- [8] Ijkel W F, Van Strien E A, Valk J M. Sequence and organization of the *Spodoptera exigua* multicapsid nucleopolyhedrovirus genome[J]. J Gen Virol, 1999, 80: 3289-304.
- [9] Ahrens C H, Russell R L Q, Funk C J, et al. The sequence of *Orygia pseudotsugata* multinuclear polyhedrosis virus genome [J]. Virology, 1997, 229: 381-399.
- [10] Haya kawa T, Ko R, Okano K, et al. Sequence analysis of the *Xestia c-nigrum* granulovirus genome [J] Virology, 1999, 262(2): 277-97.
- [11] Chen X, Li M, Sun X, et al. Genomic organization of *Helicoverpa armigera* single-nucleocapsid nucleopolyhedrovirus[J]. Arch Virol, 2001, in press.
- [12] 陈新文, 胡志红. 中国棉铃虫单粒包埋核型多角体病毒多角体蛋白基因的序列分析[J]. 中国病毒学, 1997, 12: 346-353.
- [13] Chen X, Hu Z H, Jehle J A, et al. Characterization of the ecdysteroid UDP-glucosyltransferase gene of *Helicoverpa armigera* single-nucleopolyhedrovirus[J]. Virus Genes, 1997, 15: 219-225.
- [14] Chen X, Wilfred F J, Doming C, et al. Identification, sequence analy-

- sis and phylogeny of the *lef-2* gene of *Helicoverpa armigera* single-nucleocapsid baculovirus[J]. *Virus Res*, 1999, 65: 21 - 32.
- [15] 孙修炼, 张光裕. 棉铃虫 (*Heliothis* spp) 核型多角体病毒四个分离株的比较研究[J]. *中国病毒学*, 1994, 9: 309 - 318
- [16] Ayers M D, Howard S C, Kuzio J, *et al*. The complete DNA sequence of the *Autographa californica* nuclear polyhedrosis virus[J]. *Virology*, 1994, 202: 586 - 605.
- [17] Hodgman T C. A new superfamily of replicative proteins [erratum 333, 578][J]. *Nature*, 1998, 333: 22 - 23.
- [18] Gorbalenya A E, Koonin E V, Donchenko A P, *et al*. Two related superfamilies of putative helicases involved in replication, recombination, repair and expression of DNA and RNA genomes[J]. *Nucleic Acids Research*, 1989b, 17: 4713 - 4730.
- [19] Walker J E, Saraste M, Rushwicz M J, *et al*. Distantly related sequences in the α - and β -subunits of ATP synthase, Myosin, Kinases, and other ATP-requiring enzymes and a common nucleotide binding fold[J]. *EMBO Journal*, 1988, 1: 945 - 951.
- [20] Saraste M, Sibbald P R, Wittinghofer A. The P-loop, a common motif in ATP- and GTP-binding proteins[J]. *Trends in Biological Sciences*, 1990, 15: 430 - 434.
- [21] Jacobus C, Heldens M, Yi Liu, *et al*. Characterization of a putative *Spodoptera exigua* multicapsid nucleopolyhedrovirus helicase gene[J]. *J Gen Virol*, 1997, 78: 3101 - 3114.
- [22] Liu G, Carstens E B. Site directed mutagenesis of the AcMNPV *p143* gene: effects on baculovirus DNA replication[J]. *Virology*, 1999, 253: 125 - 136.
- [23] Lohman T M, Bjornson K P. Mechanisms of helicase catalyzed DNA unwinding[J]. *Annu Rev Biochem*, 1996, 65: 169 - 214.
- [24] Matson S W, Kaiser - Rogers K A. DNA helicases[J]. *Annual Review of Biochemistry* 1990, 59: 289 - 329.