

# 大尺蠖核型多角体病毒多角体蛋白基因的结构分析\*

齐义鹏 张生家 黄永秀 裴子飞

(武汉大学病毒及分子生物学系,武汉 430072)

大尺蠖核型多角体病毒(BsNPV)的多角体蛋白基因(ocu)定位在2.3kb的BamHI-H片段上,用XhoI、SmaI、HindIII和PstI构建了H片段的物理图谱,并测定了两端636bp序列。在5'端发现了杆状病毒ocu基因共有的典型特征,即:ATG起始区;启动子区(14bp保守序列);TATA box和CATA box区等。单一XhoI位点在ATG上游-15bp处,适于作为构建ocu基因转移载体的插入位点。在这一基因5'端序列中发现了五个反转重复单元CGAGC GCTCG,讨论了这一单元在杆状病毒ocu基因高效表达中的调控功能。

**关键词** 大尺蠖核型多角体病毒;多角体蛋白基因

所有杆状病毒的多角体蛋白基因(ocu)都有一个极强大的启动子和调控序列<sup>[1]</sup>。1983年Hoofst等<sup>[2]</sup>测定了苜蓿尺蠖核型多角体病毒(AcNPV)ocu基因790bp全序列;1986年Leisy等<sup>[3]</sup>又报道了舞毒蛾NPV(OpNPV)ocu基因的全序列;在该基因5'端发现了特征性的14bp保守序列<sup>[3-6]</sup>。AcNPVocu基因转移载体的构建是从1980年Mocarski等<sup>[7]</sup>开始的,以后Smith等<sup>[8]</sup>第一次用昆虫细胞表达了 $\beta$ -干扰素。目前,表达的外源基因达到20—30种<sup>[9]</sup>,昆虫病毒基因工程得到迅速发展。

大尺蠖核型多角体病毒(Buzura suppressaria nuclear polyhedrosis virus,简称NPV)是在我国发现的感染茶树害虫的杆状病毒。我们和国内其他学者曾报道了BsNPV的一些基础性工作<sup>[10-13]</sup>,但尚未见到关于BsNPVocu基因结构和序列的报道,国外也没有BsNPV的任何文献。为了开发我国特有的这一病毒资源,使成为一个新的分子生物学模型系统,对BsNPVocu基因进行深入研究是十分必要的。

## 材料和方法

### (一) 病毒及DNA的提取

大尺蠖核型多角体病毒(BsNPV)是我们于1984年从贵州茶园分离的。取保存的毒种按文献[14]提纯病毒粒子,用TE缓冲液悬浮,加SDS到终浓度1%,65℃水浴保温30分钟,用常规酚-氯仿-异戊醇抽提,乙醇沉淀,玻棒搅起DNA,溶解在适量TE缓冲液中。

用电洗脱法从琼脂糖回收DNA片段。克隆的单链M13mp18DNA的抽提纯化方法见文献[15]。

### (二) 限制性消化和电泳

限制性核酸内切酶为Amersham和中国预防医学科学院基础医学研究所产品。按文献[16]进行单酶和双酶消化。DNA大片段用0.7%琼脂糖凝胶电泳;小片段用6%聚丙烯酰胺凝胶电泳<sup>[16]</sup>。

本文于1990年10月4日收到。

\* 国家自然科学基金和国家教委博士点基金资助的课题。

本文已在1990年澳大利亚第五届国际昆虫病毒

学术会上宣读。

### (三) 探针的制备及分子杂交

从福瑞公司购买  $a^{32}$  PdATP, 按该公司的药盒说明书制备 DNA 探针, 按文献[13]进行 Southern blot 杂交。

### (四) 分子克隆和 DNA 序列测定

M13mp18RF 载体从 Amersham 购买。用 BamHI 切开载体与 BsNPV BamHI-H 片段用 T<sub>4</sub>DNA 连接酶连接, 转化大肠杆菌 (*E. coli*) JM105 感受态细胞<sup>[16]</sup>, 在 Xgal 平板上挑取无色空斑, 经 BamHI 消化和 Southern 转膜杂交证明为阳性重组体。

从 Amersham 公司购得  $a^{32}$  PdCTP 及序列测定的其他试剂, 按该公司的实验操作和 Sanger 双脱氧链终止法进行 DNA 序列测定, 以 TBE 为缓冲液, 进行 6% 聚丙烯酰胺凝胶电泳<sup>[15,16]</sup>。

## 结 果

### (一) BsNPV 多角体蛋白基因的定 位

用含 *ocu* 基因的 AcNPV BamHI-F 片段制备放射性探针, 与 BsNPV 的 BglII、HindIII 和 BamHI 的酶切片段进行

Southern blot 转膜杂交。杂交后, 严格洗膜五次, 发现 BsNPV 的 *ocu* 基因定位在 BglII-C 和 BamHI-H 片段上 (图 1)。未发现 HindIII 片段有杂交信号, 可能是由于探针与 HindIII 的多个片段杂交, 信

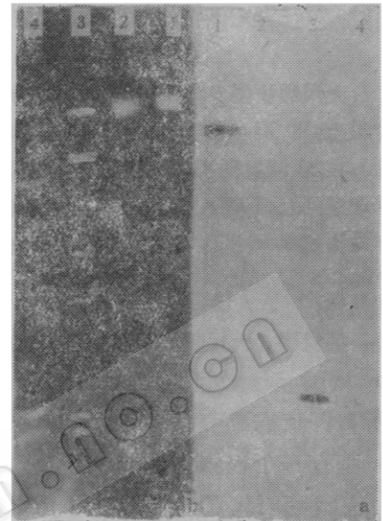


图 1 以 Southern blot 转膜杂交定位 BsNPV 的多角体蛋白基因

a. 放射自显影照片; b. 凝胶染色照片

Fig. 1 Southern blot by autoradiograph (a) and agarose gel stained by ethidium bromid (b) of BsNPV DNA digested by endonucleases

1. BglII; 2. HindIII; 3. BamHI;  
4.  $\lambda$ -DNA by HindIII as markers

表 1 H 片段的亚片段及其大小 (kb)

Fig. 1 Sub-fragments and their size from BamHI-H fragment

片 段 Fragments	XhoI	SmaI	HindIII	HindIII PstI	PstI
A	2.1	2.25	<u>2.0</u>	<u>1.4</u>	<u>1.4</u>
B	0.21*	0.08*	<u>0.35</u>	<u>0.5</u>	<u>0.55</u>
C				<u>0.35</u>	<u>0.50</u>
D				<u>0.20</u>	
合 计 Total	2.31	2.33	2.35	2.35	2.35

\* 根据序列测定结果

— 消失片段 (Fd); — 保留片段 (Fr); ~ 接头片段 (Fl)

\* Dependent on sequencing

— Disappear fragment; — Retained fragment; ~ Linker fragment

号太微弱以及由于严格洗膜所致。

我们已将 BsNPV BamHI-H 片段与质粒 pDR540 重组, 在大肠杆菌细胞中表达了多角体蛋白<sup>[2]</sup>。因此, 2.3kb 的 H 片段囊括了 *ocu* 基因的全部序列, 适于构建转移载体。

## (二) H 片段的限制性消化和物理图谱

用 HindIII、PstI、XhoI 和 SmaI 等 4 种内切酶单酶和双酶消化 H 片段, 所得片段及其大小见表 1。

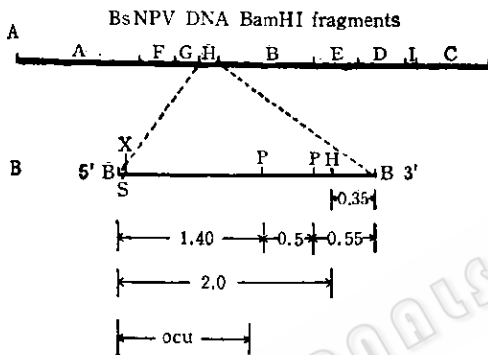


图 2 BsNPV BamHI-H 片段的物理图谱

Fig. 2 Physical map of BsNPV BamHI-H Fragment

B: BamHI; X: XhoI; S: SmaI; H: HindIII; P: PstI

根据齐义鹏等<sup>[17]</sup>提出的三三规则, 构建了这 4 个酶在 H 片段上的物理图谱 (图 2)。

在表 1 和图 2 中, SmaI 和 XhoI 的片段大小是根据序列分析结果 (详见下文) 和限制性酶切结果共同确定的。由于 HindIII-B 片段长 0.35 kb, 在所测得的

*ocu* 基因 5' 端 377bp 中未找到这一位点, 因此, HindIII-B 片段应在 3' 端, HindIII 和 PstI 的 A 片段一端为 5' 端。

## (三) H 片段的部分核苷酸序列

在以 M13mp 18 为载体克隆 BsNPV 的 BamHI-H 片段时必定有两种方向, 得到两类空斑。随机选择几个无色空斑, 测定其序列, 分别代表 H 片段的两端 (图 3), 从一个克隆测出 377bp (5' 端), 从另一个克隆测出 259bp (3' 端), 全部序列见图 3 (\*由于放射自显影微弱, 48bp 序列未能读出)。

从图 3 看出, 5' 端有杆状病毒 *ocu* 基因启动区域的典型特征, 3' 端无此特征。

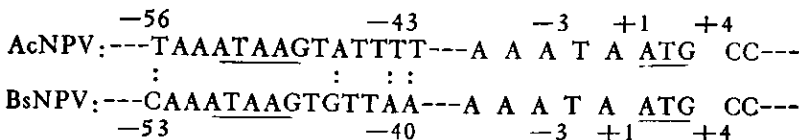
## (四) 多角体蛋白基因的结构分析

与 AcNPV *ocu* 基因比较, 在所测定的 H 片段序列中发现了杆状病毒 *ocu* 基因的一些共同特点。

首先, 杆状病毒 *ocu* 基因的 ATG 上游为 AT 富余区, 与该基因的高效表达有关<sup>[19-20]</sup>。BsNPV 和 AcNPV *ocu* 基因的这一段序列完全相同 (图 3), 均为 T<sub>A</sub><sup>C</sup> AAA-

TAATGCC, 它们都有 Kozak 结构, 即 ATG 上游 -3 为 A, 下游 +4 为 C。

其次, Rohrmann<sup>[28]</sup> 认为, 14bp 保守序列是所有杆状病毒 *ocu* 基因的共同特点, 位置从 ATG 上游 -52—-58bp 起始<sup>[4-6]</sup>。AcNPV 的这一序列从 -56bp 起, BsNPV 则从 -53bp 起。它们二者的 14 个核苷酸中仅有 4 个不同, 核心序列 ATAAG 是完全相同的。详情如下:



杆状病毒 *ocu* 基因 5' 端没有典型的 TATA box 和 CATA box<sup>[21]</sup>。我们发现,

BsNPV *ocu* 基因的非典型 TATA box (CACATTA) 位于 -81—-75 bp 间,

```

- 220
5' TGCAG CAGCT TCGCG GCGGA CAGCT GGCCA GCACT TCGGT -181
AGCTG TTGGG TGAAG ATGGC ATCTG CATA C ATCGG ATCCC -141
--
Sma
GGTA CGAGC CTCGA ATTGC TAATC ATGGT CATAG CTGTT -102
-----
I CATA box
TCCTT GTAAT TGTTA TCGCTCACATTACCACAACATACGAGCCG -58
-----
TATA box
GAAGCAAATAAGTGTTAA AGCCC TGGGG TCCCC TAATA CTA -17
-----
14bp sequence
GCTCG AGCCT CAAATAATGCCA GAC GAG CCG TAT GGA +21
-----
Xho I +1
CCC ACC GGG TTC TAC-----48bp*-----GCA AAG +90
TAT GTC CTGGACGAGCTT TGC CGG GAG TCG GAA CAA +126
-----
GCT GAG GGC GAC CCC CTA GAC TTC CGG ATA GCT +159
GTT GGG GTA GTG CTG TCA ACT ATG AAG ICT CAA +192
GCC GGT ACT GAT-----3' +204
3' GGTIG CTGGC GCTAT ATCGC GCATA CGATG GGAAG ATGGG 40
CTGCC ATTGG GCTCA TGACC GTTGT TTCGG CGTGG TATGT 80
GCAGC CCTGC GGGGA CTTTG GCGCC TACGA TACGG AGGCT 120
TACAT CTGCC CTAAT GCTGC AATAA TCGAG ACCAG GCACG 160
CTCAG ATTAT CAGCA TAACA GCAGC GGAAG GCCGA GCGCA 200
GAGTG GYCCT CCAAT TTACG CTCTA CGACT TAATT TCGAG 240
CTAGT ATATA TTGCT CATT-----5' 259

```

图 3 BsNPV 多角体蛋白基因的部分序列

↓基因起始位置; ===== CGAGC 单元

Fig. 3 Partial sequences of polyhedrin gene in BsNPV genome

↓ the gene start site, ===== the CGAGC elements

CATA box 在 -110bp 处,符合真核基因的一般特征。

因 5' 端序列中,发现了五个重复序列,其中三个在 ATG 上游,它们是:

特别值得一提的是在 BsNPV ocu 基

-136

-69

-13

GTA CGAGC CT----ATA CGAGC CG----GCT CGAGC CT  
CAT GCTCG GA----TAT GCTCG GC----CGA GCTCG GA

这五个重复序列均以反转重复单元 CGAGC GCTCG 为核心。十分巧合, 在 AcNPV *ocu* 基因<sup>[2]</sup> 5' 端的 -198 bp 和 OpNPV *ocu* 基因<sup>[4]</sup> 5' 端的 -170bp 和 -131bp 处也发现了这一核心单元。估计, 它们对杆状病毒 *ocu* 基因的高效表达有某种特殊功能。

## 讨 论

在 BsNPV、AcNPV 和 OpNPV 的 *ocu* 基因 5' 端都发现有 CGAGC GCTCG 反转重复单元。Guarino 等指出<sup>[19]</sup>, 基因的重复

序列一般都有某种特殊功能。

例如, AcNPV 基因组中的五个同源重复序列 (hr1—hr5) 有增强子功能<sup>[20]</sup>, 恰好, 在这五个增强子中找到了大量的 GCTC 和 ACTC 单元。最近 Dennis 等报道<sup>[22]</sup>, 在  $\alpha$ -1 抗胰蛋白酶基因 ( $\alpha$ 1-AT) 和转 thyretin 基因 (TTR) 的启动子上游各有一个增强子, 它们分别包含有以 ACTC 为核心单元的重复序列。进一步发现, ACTC 是细胞中顺式作用蛋白的结合位点, 赋予了增强子增强基因表达的功能。他们将 ACTC 突变, 增强子活性随之消失。因此, 我们将本文发现的 GCTCG 序列同上面引用的资料进行比较, 即:

BsNPV	<i>ocu</i>	1	CCATGCTCG	OpNPV	<i>ocu</i>	2	TGTTACTCG
		2	GTATGCTCG	AcNPV	hr5	1	TTCAACTCG
AcNPV	<i>ocu</i>	1	AGCCGCTCG			7	TATGACTCA
OpNPV	<i>ocu</i>	1	CAGGGCTCT			8	TTCTACTCG
AcNPV	hr5	2	AAATGCTCA			9	CTGAACTCG
		3	ATTGGCTCG			10	TTTAACTCG
		4	AAAIGCTCA			11	CTAAACTCG
		5	GCTAGCTCA	$\alpha$ 1-AT	A		ACCTACTCT
		6	AAAIGCTCA		B		TGIGACTCA
					C		TAAGACTCC
				TTR	1		TAAIACTCT
					3		ICTIACTCA

众所周知, 杆状病毒 *ocu* 基因是一个表达水平非常高的基因, 长期以来, 其高效表达的原因得不到合理解释, 通过上述分析, 我们假设在增强子中有两类功能单元, 即 ACTCA 型和 GCTCG 型, 杆状病毒 5' 端非编码区很可能存在着增强子序列, 其中的 GCTCG 单元可能是多角体蛋白基因得以高效表达的原因之一。

## 参 考 文 献

- [1] Smith, G. E. et al.: *J. Virology*, **45**: 215, 1983.
- [2] Hooft, B. J. et al.: *Virology*, **131**: 561, 1983.
- [3] Leisy, D. et al.: *Virology*, **153**: 280, 1986.
- [4] Leisy, D. et al.: *Virology*, **153**: 157, 1986.
- [5] Wilson, M. F. et al.: *J. Virology*, **61**(3): 661, 1987.
- [6] Possee, R. et al.: *Nucl. Acid Res.*, **15**: 10233, 1987.

- [7] Mocarski, E. S. et al.: *Cell*, **22**:243, 1980.
- [8] Smith, G. E. et al.: *Mol. Cell Biology*, **3**(12): 2156, 1983.
- [9] 齐义鹏: 病毒学报, **6**(4): 383—390, 1990.
- [10] 齐义鹏等: 病毒学报, **3**(3): 264, 1987.
- [11] 黄永秀等: 武汉大学学报, **1**: 121, 1988.
- [12] 齐义鹏等: 微生物学报, **31**(1): 25—31, 1991.
- [13] 梁布锋等: 病毒学杂志, **4**: 402, 1989.
- [14] Maskos, C. B. et al.: *J. Invert. Pathology*, **37**:174, 1981.
- [15] 齐义鹏: 中国科学(B辑)**19**: 609, 1986.
- [16] 齐义鹏等: 基因工程原理和方法, 第 14 页, 329 页, 335 页, 203 页, 四川大学出版社, 成都, 1989 年.
- [17] 齐义鹏等: 武汉大学学报, **2**: 89—96, 1990.
- [18] Rohrmann, G. F.: *J. Gen. Virology*, **67**: 1499, 1986.
- [19] Guarino, L. et al.: *J. Virology*, **60**(1):215, 1986.
- [20] Guarino, L. et al.: *J. Virology*, **60**(1): 224, 1986.
- [21] Qin, J.-C. et al.: *J. Gen. Virology*, **70**: 1273, 1989.
- [22] Dennis, R. G. et al.: *Science*, **239**:786, 1988.

## STRUCTURAL ANALYSIS OF POLYHEDRIN GENE IN BUZURA SUPPERISSARIA NUCLEAR POLYHEDROSIS VIRUS GENOME

Qi Yipeng Zhang Shengjia Huang Yongxiu Pei Zifei

(Department of Virology, Wuhan University, Wuhan 430072)

The polyhedrin gene (ocu) of Buzura superpissaria nuclear polyhedrosis virus (Bs-NPV) was mapped on BamHI-H fragment were sequenced by Sanger ddNTP chain-fragment was constructed by XhoI, SmaI, HindIII, PstI and its 636 bp on both ends were sequenced by Sanger ddNTP chain-termination method. The typical properties of baculovirus ocu genes are found on 5' end of Bs-NPV ocu gene, which are ATG start region, 14 bp conservative sequences, TATA

box and CATA box etc. The unique XhoI site, which will be inserting site in transfer vector constructed by Bs-NPV ocu gene, is mapped on-15 bp upstream of ATG. Specific 5 boxes that are CGAGC sequences are observed on the gene. The important function in high expression of the gene is discussed.

### Key words

Buzura superpissaria nuclear polyhedrosis virus; Polyhedrin gene