

# 文山松毛虫质型多角体病毒(DpwCPV)NS5 蛋白基因的 cDNA 克隆及序列分析

王 琼 张珈敏 文 力 杨 娟 胡远扬\*

( 武汉大学病毒研究所 武汉 430072 )

**摘 要** 通过对文山松毛虫质型多角体病毒(*Dendrolimus punctatus* Wenshanensis cytoplasmic polyhedrosis virus, DpwCPV)的增殖、纯化,获得一株单一类型的质型多角体病毒。提纯的病毒粒子经 SDS-热酚法抽提得到基因组 dsRNA,使用低熔点琼脂糖凝胶电泳分离并回收纯化第九片段 S9。S9 RNA 双链经高温变性,逆转录合成 cDNA 双链。根据 DpwCPV 与 BmCPV-1 的同源性设计引物,将 S9 进行 PCR 扩增后,克隆到 PMD18-T 载体上。最终获得一个 977bp 的序列,其中包含一个 963bp 的开放阅读框(ORF)。推测 DpwCPV S9 基因编码一个 320 个氨基酸的蛋白,分子量约为 35560。

**关键词**:文山松毛虫 文山松毛虫质型多角体病毒 NS5 蛋白基因

中图分类号:Q934.2 文献标识码:A 文章编号:1001-6209(2003)01-0021-08

文山松毛虫(*Dendrolimus punctatus* Wenshanensis T.)是我国主要的林业害虫,每年造成巨大的林业资源损失和经济损失。文山松毛虫质型多角体病毒(*Dendrolimus punctatus* Wenshanensis cytoplasmic polyhedrosis virus, DpwCPV)是我国特有的一种质型多角体病毒<sup>[1]</sup>,在森林生物防治方面可以发挥重要的作用<sup>[2]</sup>。此病毒 1993 年首先由云南林业科学研究所和武汉大学共同分离鉴定。其基因组由十个节段的双链 RNA 组成,但在 1% 琼脂糖凝胶电泳时只呈现 9 条带,其中第二条(约 3850bp)和第三条(3840bp)无法分开。其电泳图谱与家蚕质型多角体病毒 1 型(*Bombyx mori* Cytoplasmic Polyhedrosis Virus Type 1, BmCPV-1)一致,推测它们同属 1 型 CPV。

目前对于 DpwCPV 的研究只限于核酸分离、组织病理、免疫学、复制方式、提纯及加工工艺、安全实验及大面积林间防治等方面,其分子生物学水平的研究报道较少<sup>[3-5]</sup>。其所属的 CPV 1 型病毒目前也只有 BmCPV-1 的两个毒株(H 株和 I 株)以及一种舞毒蛾质型多角体 1 型(*Lymantria dispar* Cytoplasmic Polyhedrosis Virus Type-1, LdCPV-1)的全序列得到了测定。这与呼肠孤病毒其他属的研究相比,明显滞后。通过对 DpwCPV 的各个基因组片段 cDNA 克隆,构建其 cDNA 文库,然后进行各基因组片段序列的测定,不仅可以填补此研究领域的空白,为该病毒的同源性研究、分类研究、进化研究、蛋白质空间结构研究奠定基础,同时也为其应用研究提供了理论依据。本实验室对 DpwCPV 第 9 片段 S9 的编码序列

\* 通讯作者, E-mail: yyhu@whu.edu.cn

作者简介:王 琼(1977-),女,现从事昆虫病毒理论及应用研究

收稿日期:2002-06-06,修回日期:2002-09-09

进行了 cDNA 克隆和序列测定。通过序列分析比较发现, DpwCPV 和 BmCPV 第九片段的编码区存在 136 个核苷酸的差异, 均为碱基转换, 其中 114 个是同义突变, 体现在氨基酸水平上有 22 个氨基酸的变化。DpwCPV 和 LdCPV 第九片段的编码区存在 158 个核苷酸的差异, 均为碱基转换, 其中 126 个是同义突变, 体现在氨基酸水平上有 32 个氨基酸的变化。现将结果报告如下。

## 1 材料和方法

### 1 材料

**1.1.1 病毒:** 文山松毛虫质型多角体病毒( DpwCPV )由本实验室保藏。

**1.1.2 试剂:** 逆转录酶 M-MLV 及 dNTP 购自 Promega 公司; 限制性内切酶、高保真 Taq<sup>ex</sup> 酶和 pMD18-T 载体( 2692bp )试剂盒购自 TaKaRa 公司; *E. coli* DH5 $\alpha$  由本实验室保存; RT-PCR 引物由上海生工合成, 序列如下:

引物 1 5' - GATCATCATGGAAGCCTTCTTACTC - 3'

引物 2 5' - AGCGACACTATTCTGATACATC - 3'

引物 3 5' - GACGATAGGCTGCACCTTC - 3'

### 1.2 方法

**1.2.1 病毒的增殖、纯化和 dsRNA 的抽提:** 病毒的增殖、纯化参考 Arella 等人的方法<sup>[6]</sup>, 经过差速离心、蔗糖梯度离心和多角体的碱解, 获得纯化的 DpwCPV 江西株病毒粒子。采用 SDS-热酚法从病毒粒子中纯化 RNA。

**1.2.2 S9 的分离纯化:** 江西株 CPV 核酸经 1% 的低熔点琼脂糖凝胶电泳分离, 于紫外灯下切下含 S9 的凝胶, 采用低熔点琼脂糖法回收。经无水乙醇沉淀后, 溶于少量无 RNA 酶的 ddH<sub>2</sub>O 中。

**1.2.3 cDNA 第一链的合成及其 PCR 扩增:** 取 500ng 纯化的基因组单一片段 S9 溶于无 RNA 酶的 ddH<sub>2</sub>O 中, 100℃ 水浴加热 5min 使之变性。cDNA 第一链合成参考 Hagiwara 等人的方法<sup>[9]</sup> 取 500ng RNA 样品与引物 1 和引物 2 各 50pmol, 在 70℃ 下退火, 然后再加入 dNTP, M-MLV 逆转录酶, 合成 cDNA 第一条链。随后进行 PCR 扩增, 加入引物 1 和 2 各 50pmol, Taq<sup>ex</sup> 聚合酶 2.5 单位, 扩增条件: 94℃ 变性 30s, 52℃ 退火 90s, 72℃ 合成 60s, 30 个循环后, 于 72℃ 放置 10min 使延伸反应充分进行。由此得到 405bp 片段。

使用引物 1 和引物 3 作为 RT-PCR 引物, PCR 退火温度为 50℃, 其他条件不变, 由此得到 679bp 片段。

**1.2.4 cDNA 克隆:** 用低熔点琼脂糖凝胶电泳回收纯化 PCR 产物。按照 pMD18-T 载体系统操作手册, 回收的 PCR 产物和载体 DNA 的量分别为 300ng 和 100ng, 连接过夜。连接产物转化 *E. coli* DH5 $\alpha$  感受态细胞后, 涂于含有 Amp/IPTG/X-gal 的平板。

**1.2.5 重组质粒的筛选和鉴定:** 经选择培养基筛选, 碱裂解法快速制备质粒 DNA, 琼脂糖凝胶电泳及酶切鉴定重组 DNA。

**1.2.6 DNA 序列测定和分析:** 以 cDNA 阳性克隆子作模板, 引物采用质粒载体上的通用引物, 正、反向进行双链测序。自动测序由上海博亚公司完成。

## 2 结果

### 2.1 病毒的 dsRNA 图谱

DpwCPV-dsRNA 在 1% 的琼脂糖凝胶电泳结果见图 1。此核酸图谱显示,文山松毛虫 CPV 的核酸基因组至少有 10 个片段,与舞毒蛾质型多角体病毒和家蚕质型多角体病毒 1 型(BmCPV-1)基因组图谱一致。

### 2.2 S9 的回收和 RT-PCR

图 1 中可看到,回收的 S9 的迁移率与基因组中对应片段完全吻合。根据已发表的 BmCPV-1 S9 的全序列<sup>[7]</sup>,设计合成包含起始密码的上游引物 1 和终止密码的下游引物 2。RT-PCR 扩增结果见图 2。扩增产物约为 400 bp,经过测序,该片段长 405 bp,与 BmCPV 的 S9 序列 660 ~ 1050 bp 高度同源,且两端序列均与引物 2 匹配,推测是下游引物 2 在 660 bp 左右处反向错配,阻断了预计的全长编码序列的合成。根据已经得到的序列设计引物 3,使用引物 1 和引物 3 进行 RT-PCR 反应结果见图 3。产物大小略大于 650 bp,符合预期的结果,经过测序,该片段长 679 bp,与 BmCPV 的 S9 序列 70 ~ 750 bp 高度同源。



图 1 DpwCPV 基因组及第九片段的回收

Fig.1 DpwCPV genome and separation and purification of S9

1. Marker  $\lambda$ DNA/*Eco*RI + *Hind*III;
2. DpwCPV genome 3. Purified S9.

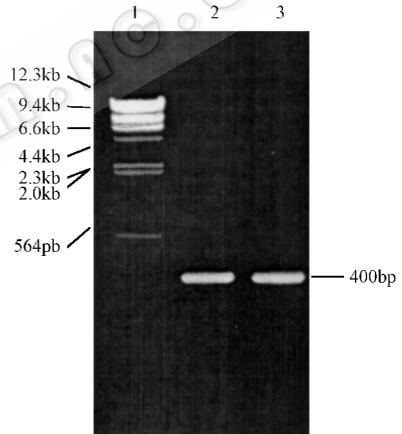


图 2 DpwCPV 的 S9 片段的 PCR 扩增

(引物 1 和 2)与阳性克隆鉴定

Fig.2 PCR of S9 and PCR analysis of recombination plasmid (pMD18T/400bp)

1. Marker  $\lambda$ DNA/*Hind*III 2. PCR of S9 (Primer 1 and 2);
3. PCR of insert segment in recombination plasmid (pMD18T/400bp).

### 2.3 重组质粒的酶切鉴定

小 PCR 片段的重组子限制性内切酶鉴定结果见图 4。阳性重组质粒经 *Eco*RI/*Hind*III 双酶切后,得到大小两个不同的片段。较大的片段是载体质粒,约为 2690 bp,与载体大小相符,较小片段是插入片段,约为 400 bp,与 DpwCPV 第 9 片段的较小的 PCR 产物的大小相符。

大 PCR 片段的重组子限制性内切酶鉴定结果见图 5。阳性重组质粒经 *Eco*RI/*Hind*III 双

酶切后,得到大小两个不同的片段。较大的片段是载体质粒,约为 2690 bp,与载体大小相符。较小片段是插入片段,约为 650 bp, DpwCPV 的第 9 片段的较大的 PCR 产物的大小相符。

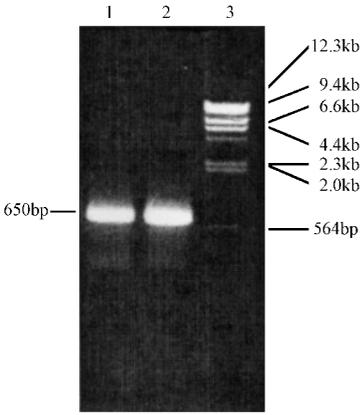


图 3 DpwCPV 的 S9 片段的 PCR 扩增 (引物 1 和 3)与阳性克隆鉴定

Fig.3 PCR of S9 and PCR analysis of recombination plasmid ( pMD18T/650bp )

1. Marker  $\lambda$ DNA/ *Hind* III ;
2. PCR of S9 (Primer 1 and 3) ;
3. PCR of insert segment in recombination plasmid ( pMD18T/650bp ).

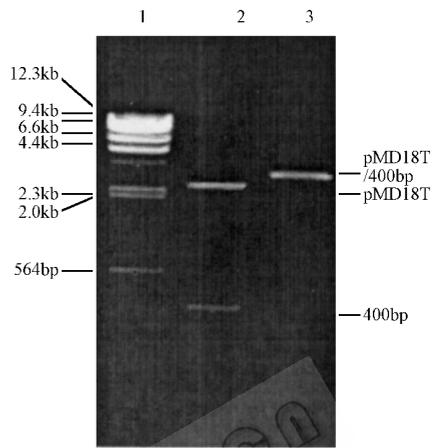


图 4 重组质粒的限制性酶切鉴定

Fig.4 Restriction analysis of recombination plasmid ( pMD18T/400bp )

1. Marker  $\lambda$ DNA/*Hind* III ;
2. Restriction analysis of recombination plasmid ( pMD18T/400bp ) by *Eco*RI and *Hind* III ;
3. Recombination plasmid ( pMD18T/400bp ).

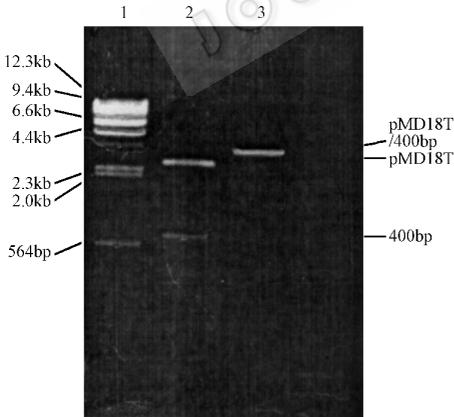


图 5 重组质粒的限制性酶切鉴定

Fig.5 Restriction analysis of recombinant plasmid ( pMD18T/650bp )

1. Marker  $\lambda$ DNA/*Hind* III ;
2. Restriction analysis of recombination plasmid ( pMD18T/400bp ) by *Eco*RI and *Hind* III ;
3. Recombination plasmid ( pMD18T/650bp ).

### 2.4 重组质粒的 PCR 鉴定

图 2 中第 1 泳道是以小 PCR 片段的阳性克隆子为模板,用载体的通用引物对插入片段进行 PCR 的结果。其大小与第 2 泳道中 DpwCPV 的第 9 片段的较小的 PCR 产物的大小相符。

图 3 中第 1 泳道是以大 PCR 片段的阳性克隆子为模板,用载体的通用引物对插入片段进行 PCR 的结果。其大小与第 2 泳道中 DpwCPV 的第 9 片段的较大的 PCR 产物的大小相符。

### 2.5 序列测定和分析

S9 RT-PCR 产物克隆片段的序列测定结果如图 6。

利用大片段 3'端和小片段 5'端之间 89bp 完全匹配的碱基序列,可拼接出一个 977bp 的序列 (405bp 片段 5'端的 16 个不与 679bp 片段匹配的碱基是错配的引物序列),其中包含一个 963bp

大小的 ORF 起始子位于第 8 ~ 10 位碱基,终止子位于第 968 ~ 970 位碱基。此片段的 G + C 含量为 43.51% , $T_m$  值为 82.22°C。推测文山松毛虫质型多角体病毒基因组的第 9 片段编码一种由 320 个氨基酸的蛋白,分子量是 35560。其中有 32 个强碱性氨基酸(K,R);31 个强酸性氨基酸(D,E);107 个疏水性氨基酸(A,I,L,F,W,V);102 个极性氨基酸(N,C,O,S,T,Y,X)。多肽的等电点是 7.721,在 pH7.0 时,带 2.241 个单位的正电荷。

A:405bp 片段序列:

```
AGCGACACTA TACTGATACA TCACCACCCC CTATCCAATT AATACAATAC GGTGTGACGC 60
CATTTCATATC AGCAGAAATC ATCTGTGAAG GTGCAGCCTA TCGTCCAATC GATGTTGACA 120
ATTACGCAAT GAATAGTAAT CATCTAGTGT CATATGCTCC GTTCTTCGTA CCAAACGATA 180
CTAAGCCTGG AAGCCGTATA GATTtGTTGA TGGTCAACCA TCTAAAAAAG TTCAATTTGG 240
TATTCGATTC TAGGTACAAA ACTGGTGGCA GTGTAATGGT ATCTTCTCGT CCCGAGCGTG 300
CACAAAACGA GACTTCAGCC AGTCAAATCA TGCCGACGTC TGTGAAGCAT ATAGCTAATG 360
ATGATTTAAA TGTAGATGAT GGTGATGTAT CAGAATAGTG TCGET 405
```

B:679bp 片段序列:

```
TGCCTGTAG GTCGACGATT GATCATCATG GAAGCCTTCT TACTCGAAAA CCGCAAACCC 60
AAAATCACC A CTCTGCATC AGGAAAGACT CTCAAACCAG CGACGCATCG TTTAAATTTG 120
CCAGCATACA CGAAGTTGAT TCATGAAC TGAGGACCAAGA CGCAGCTAA AATTACGATT 180
TCCCTCTCAA CCGAAAGTCA GATACATATG GTATGGGTAA AATCAGGTCT TGTTTTCTTT 240
ACGCCCAGCG CATTCACCC AGCATATGTA AATTTTCGAT CCTCGAACAG ACTTACCGAT 300
GTGCCAGCAC TCACCAAATC TACGTTCCCG CAATCCGACG TGAAGCTTGT AGAGTCTACT 360
CCCCTGCCTA ATGATGAGT TTCACATGTC GCGTCAATTC AACTCATCAT GTGGAAAGAT 420
GGAGCGTTAA GCGTTCTGAA TGATCTGAGC AAATGTGCGA TCACCTTCAT CAATCAATGC 480
GAGGATACGT TCAAAAGTGG CACCAATTTA AACAAGGAGA TGTACAATAG ATGTATCACG 540
GCAGAGTCTC GCGACTTCTG CAACCAGATG AAACCTCGTC TGATTGGACG TCTATGTTAC 600
GGACAAACTA CATCACCACC CCCTATCCAA TTATATCAAT ACGGTGTGAC GCCATTCATA 660
TCAGCAGAAA TCATCTGTG 679
```

图 6 DpwCPV S9 的 RT-PCR 产物的序列测定

Fig.6 The sequencing of segment 9 in DpwCPV

A 405bp ; B 679bp.

图 7 将文山松毛虫质型多角体病毒基因组的第 9 片段与 BmCPV 和 LdCPV 基因组的第 9 片段编码区进行了比较,结果图 7 所示。DpwCPV 和 BmCPV-1(I 株)的 S9 编码区序列相比较,核苷酸序列同源性的 85%,二者有 136 个核苷酸的差异,其中 114 个出现在密码子的第三位。DpwCPV 和 LdCPV 的 S9 编码区序列相比较,核苷酸序列同源性的 83%,二者有 158 个核苷酸的差异,其中 126 个出现在密码子的第三位。

图 8 对 DpwCPV、LdCPV 和 BmCPV-1(I 株)的基因组第 9 片段编码的蛋白质(NS5 蛋白)进行了比较。如图 8 所示,DpwCPV 第九片段推导的蛋白氨基酸序列和 BmCPV-1(I 株)相比较,同源性的 93%,二者有 22 个氨基酸的差异。DpwCPV 与 LdCPV 比较,同源性为 89%,二者有 32 个氨基酸的差异。经 DNASTAR 分析,发现这些变化对多肽的亲水性,疏水性没有显著影响。DpwCPV 与 BmCPV 间 22 个氨基酸的变化中只有位于第 98,203,272 和 273 位的 4 个引起了蛋白质疏水特性的改变。DpwCPV 与 LdCPV 比较,32 个氨基酸转换中只有 13 个影响到蛋白质的疏水性。它们是位于第 127,130,131,135,136,203,273,289,293,295,309,317,318 位的氨基酸变化。除此之外,三种 CPV 的 NS5 蛋白均表现为疏水性,这与 NS5 蛋白在受染细胞核内出现,参与病毒核酸的复制有密切联系。但是这些氨基酸的差异,对蛋白质二级结构有一定影响,而且主要体现 NS5 蛋白 C 端。例如,DpwCPV 与 BmCPV 比较,第 203 位氨基酸的变化使 BmCPV 在 203 ~ 206 比 DpwCPV 多出一个

	***								
DpwCPV	GATCATCATG	GAAGCCTTCT	TACTCGAAAA	CCGCAAACCC	AAAATCACCA	CTCTCGCATC	AGGAAAGACT		70
LdCPV	.....	.....	.....	.....	C.....	.....	.....		
BmCPV	.....	.....	.....	.....	.....T.....	.....	.....		
DpwCPV	CTCAAACCAG	CGAGCGATCG	TTTAAATTTG	CCAGCATAACA	CGAAGTTGAT	TCATGAACTG	AGGACCAAGA		140
LdCPV	.....	...A.....	...G·C...	...G.....	.....	.....T·A	C...A...		
BmCPV	.....	.....C..	...G.....A	·G.....	A.....	.....A	C.....A.		
DpwCPV	CGCACGCTAA	AATTACGATT	TCCCTCTCAA	CCGAAAGTCA	GATACATATG	GTATGGGTAA	AATCAGGTCT		210
LdCPV	...T·C·	GG·GG·C·	·G·T·	·T·T·C·	A.....	.....	.....		
BmCPV	.....C·	·G.....	·T·T·	...G·C·	.....	.....	.....		
DpwCPV	TGTTTTCTTT	ACGCCAGCG	CATCTCACCC	AGCATATGTA	AATTCGCAT	CCTCGAACAG	ACTTACCGAT		280
LdCPV	.....	.....T.....	·G.....T·	·G.....G	.....	.....	.....		
BmCPV	...C.....C	.....T...	·G.....T·	·G.....G	.....	.....G...	.....T...		
DpwCPV	GTGCCAGCAC	TCACCAAATC	TACGTTCCCG	CAATCCGACG	TGAAGCTTGT	AGAGTCTACT	CCCCTGCCTA		350
Ldcpv	.....	.....T·A·	A.....	.....A·T·	.....T·A·	.....A·G·	.....T.....		
BmCPV	.....	.....C·	G.....	.....T·	C...T·AA·	C.....	...T·A...		
DpwCPV	ATGATGAGAT	TTCACATGTC	GCGTCATTTT	AACTCATCAC	GTGGAAAGAT	GGAGCGTTAA	GCGTTCTGAA		420
LdCPV	·C...A·C	C.....	·A...C·	·G·G·A·	T...TG·G	...AC·G·	A·A.....		
BmCPV	...C...GC	·T.....	·A.....	...G·T·	A.....	.....	.....		
DpwCPV	TGATCTGAGC	AAATGTGCGA	TCACTTTCAT	CAATCAATGC	GAGGATACGT	TCAAAGTGG	CACCAATTTA		490
LdCPV	...T...T	·G.....	·G.....	.....	.....	.....	A·T...C·		
BmCPV	...C...C	·G.....A·	·T·G.....	T...G·T	.....A·	.....	...G·C·G		
DpwCPV	AACAAGGAGA	TGTACAATAG	ATGTATCAGC	GCAGAGTCTC	GCGACTTCTG	CAACCAGATG	AAACTCGTGC		560
LdCPV	.....T	.....C·	...C·T...	.....G·	T.....	.....	...GT·T·AT		
BmCPV	.....	.....	.....T·T	.....	...T.....	...T.....	...T...A·		
DpwCPV	TGATTGGACG	TCTATGTTAC	GGACAAACTA	CATCACCACC	CCCTATCCAA	TTAATACAAT	ACGGTGTGAC		630
LdCPV	·A·C.....	.....T	·C·G·A·	·G...G·	T·G.....	·GTAT·	...C.....		
BmCPV	...C.....	·T·G.....T	·G.....	.....	T...T·	·TAT·	.....A·		
DpwCPV	GCCATTTCATA	TCAGCAGAAA	TCATCTGTGA	AGGTGCAGCC	TATCGTCCAA	TCGATGTTGA	CAATTACGCA		700
LdCPV	A.....C	·T...C·	...T...C·	...T·T	.....T·	T.....	G...T·C		
BmCPV	A...T...	...G·C·	...C·	...C·C·T	.....T...	·T...C·	A·C·T...		
DpwCPV	ATGAATAGTA	ATCATCTAGT	GTCATATGCT	CCGTTCTTCG	TACCAAACGA	TACTAAGCCT	GGAAGCCGTA		770
LdCPV	...C...	...CAGC·	A·T...C	·A...T·	...T·T·	G·C...G	·T·T·C·		
BmCPV	...C...	.....G·	A·C·C·	·A.....	·G...T·	·C·A·C	.....C·		
DpwCPV	TAGATTTGTT	GATGGTCAAC	CATCTAAAAA	AGTTCAATTT	GGTATTTCGAT	TCTAGGTACA	AAACTGGTGG		840
LdCPV	·C·C·AA·	A.....T...	.....T.....	·A...C·	·A·G.....	·T.....	.....		
BmCPV	·T...A·	.....T...	.....C·G·	·A·T·CC·	AA.....	A·GT.....	...C.....		
DpwCPV	CAGTGTAAATG	GTATCTTCTC	GTCCCAGACG	TGCACAAAAC	GAGACTTCAG	CCAGTCAAAAT	CATGCCGACG		910
LdCPV	T...G...	.....T.....	...T...A...	·A·G·T	·G.....	TT·C·G·	T...T·T		
BmCPV	G·C·G...	·G·C...	.....A...	·G.....	·AG...C·	·A·A...	.....A...		
DpwCPV	TCTGTGAAGC	ATATAGCTAA	TGATGATTTA	AATGTAGATG	ATGGTGATGT	ATCAGAATAG	TGTCGCT		977
LdCPV	...A·A·	...T.....	...A.....	·CG·C·	·C...A·G	.....G...	.....		
BmCPV	·TG·T...	...T·G·	...C·C·	G·A·G...	.....	.....	.....		

图 7 DpwCPV 与 BmCPV-1(I 株)和 LdCPV 的 S9 编码区序列的比较

Fig. 7 Comparison of DpwCPV S9 with that of BmCPV-1 I strain and LdCPV

Identical nucleotides are indicated by “.”;

The initiation codon and termination codon are indicated by “\*\*\*”.

Turn 结构;DpwCPV 与 LdCPV 比较,第 49,50 位氨基酸的变化使 DpwCPV 在 47~53 位的 β-折叠在 LdCPV 中成为 α-螺旋。

### 3 讨论

DpwCPV	MEAFLENRK	PKITTLASGK	TLKPATHLRN	LPAYTKLIHE	LRTKTHAKIT	ISLSTESQIH	MVWVKSGLVF	70
LdCPV	.....	·Q·.....	.....	.....	.....	.....VA	.....D.....	.....
BmCPV	.....	.....	.....	.....	.....	.....V·	.....	.....
DpwCPV	FTPSASHPAY	VNFASSNRLT	DVPALTKSTF	PQSDVKLVES	TPLPNDIEISH	VASFQLITWK	DGALSVLNDL	140
LdCPV	.....	.....	.....T·	.....I·T	.....	.....V·M	E·NI·	.....
BmCPV	.....	.....S·	.....P·	.....I·	.....A·	.....V·	.....I·	.....
DpwCPV	SKCAITFINQ	CEDTFKSGTN	LNKEMYNRCI	TAESRDFCNQ	MKLVLIIGRLC	YGQTTSPPI	QLIQYGVTFP	210
LdCPV	.....S·	.....	.....L·	.....	.....F·	.....	.....Y·	.....
BmCPV	.....S·	.....	.....	.....	.....F·	.....	.....Y·	.....
DpwCPV	ISAEIICEGA	AYRPIDVDNY	AMNSNHLVSY	APFFVPNDTK	PGSRIDLMMV	NHLKKNLVF	DSRYKTGGSV	280
LdCPV	·SD·	.....E·	.....T·	.....E·	.....I·	.....M·	·W·	.....
BmCPV	·D·	.....S·E·	.....	.....	.....	.....I·	·TW·	.....
DpwCPV	MVSSRPERAQ	NETSASQIMP	TSVKHIANDD	LNVDGQDQVSE				320
LdCPV	.....T·	·A·V·	.....E·	·TA·EG·				
BmCPV	.....	·A·N·	ML·	·DM·				

图 8 DpwCPV 和 BmCPV-1(I 株),LdCPV 第九片段编码的蛋白质的氨基酸序列比较

Fig.8 Comparison of deduced amino acid sequence of DpwCPV S9 with that of BmCPV-1 I strain( Identical amino acid is indicated by “.” )

1 型 CPV。在 CPV 的 14 个电泳类型中<sup>[8]</sup>,目前只有三种 1 型 CPV ,BmCPV-1 的 I 株、H 株 ,LdCPV-1 和一种未知型 CPV ,TnCPV 的第 9 核酸片段( S9)的序列被测定。通过比较发现 ,DpwCPV 和三种 1 型 CPV 间核酸序列同源性均在 83% 以上 ,它们编码的蛋白均由 320 个氨基酸组成 ,氨基酸序列同源性在 89% 以上。而在不同型的 CPV 第九片段的序列间 ,几乎找不到同源性。TnCPV S9 编码的蛋白由 285 个氨基酸组成 ,在大小上就与 1 型 CPV 的 NS5 蛋白不同。此前 ,笔者对 DpwCPV 的多角体蛋白基因进行了序列测定 ,并与各种已知的 CPV 多角体蛋白基因进行了比较分析 ,发现 DpwCPV 与 1 型 CPV 在此基因上的同源性达到 88% 以上 ,而此基因编码的蛋白质的同源性高达 97% 以上。显然 ,相对于编码多角体蛋白的 S10 片段来说 ,S9 片段的变异性更大 ,但在同型 CPV 之间仍具有相当高的保守性。这些结果为 DpwCPV 属于 1 型 CPV 的推论提供了证据。

对于三种 CPV NS5 蛋白 N 端的 120 个氨基酸来说 ,它们不仅有高度一致的相似的疏水性区域和亲水性区域 ,而且显示相似的二级结构 ,这些区域的二级结构富含  $\alpha$ -螺旋和转角结构。DpwCPV 的 NS5 的 N-末端有 3 个不同的基序区 :Arg-Lys-X-Lys ,Lys-XX-Lys 和 Arg-X-Lys。他们分别位于 9→12 ,20→23 ,42→44。这几个基序同样存在于 BmCPV 的 NS5 蛋白中 ,它们与呼肠病毒另一成员人轮状病毒 NS26 蛋白的 C-末端具有 46% 的同源性 ,它们也同样存在于牛轮状病毒 NS26 蛋白的 3'-非翻译区。NS5 蛋白 N 端的保守性意味着 ,它们可能在决定 S9 片段编码蛋白的功能方面有决定性作用 ,是调节 CPV 核酸复制所必需的。

与 BmCPV 的 NS5 蛋白一样 ,DpwCPV 的 NS5 不含 GDD。GDD 作为 NTP-结合位点 ,在大多数 RNA-依赖的 RNA 聚合酶中是保守的。对 BmCPV 的分析显示 ,其第九片段编码的 NS5 蛋白是一种非结构蛋白 ,只在病毒复制的早期表达 ,并能与 dsRNA 基因组结合 ,可能是一种基因表达的调控蛋白。由其 ORF 预测的 NS5 蛋白序列中 ,第 36~43 位碱基的序列 KLIHELRT 与人脊髓灰质炎病毒( Human Poliovirus )的 RNA 复制酶( RNA-dependent RNA polymerase )的基序( motif )YVKDELRS 有 63% 的同源性<sup>[7]</sup> ,而这一段序列在 BmCPV-1

(I 株和 H 株) LdCPV-1 中均是保守的。据此推测,该序列所在的区域可能是 NS5 蛋白的重要结构域。

### 参 考 文 献

- [ 1 ] 陈昌杰. 松毛虫综合管理. 北京: 林业出版社, 1990. 282 ~ 300.
- [ 2 ] 梁东瑞, 胡远扬. 松林害虫病毒生物防治工程. 武汉: 武汉大学出版社, 1986. 17 ~ 19
- [ 3 ] Echeverry F, Bergeron J, Kaupp W, et al. Sequence analysis and expression of the polyhedrin gene of *Choristoneura fumiferana* cytoplasmic polyhedrosis virus ( CfCPV ). *Gene*, 1997, **198**: 399 ~ 406.
- [ 4 ] Galinski M S, Yu Y, Hemingway B R, et al. Analysis of the C-polyhedrin genes from different geographical isolates of a type 5 cytoplasmic polyhedrosis virus. *J Gen Virol*, 1994, **75**: 1969 ~ 1974.
- [ 5 ] Hagiwara K. Nucleotide sequences of *Bombyx mori* cytoplasmic polyhedrosis virus segments 6 and 7 encoding viral structural protein V4 and V5. *J Gen Virol*, 2000, **81**: 1143 ~ 1147.
- [ 6 ] Arella M, Lavallee C, Belloncik S, et al. Molecular cloning and characterization of cytoplasmic polyhedrosis virus polyhedrin and a viable deletion mutant gene. *J Virol*, 1988, **62**: 211 ~ 217.
- [ 7 ] Hagiwara K, Tomita M, Nakai K, et al. Determination of the nucleotide sequence of *Bombyx mori* cytoplasmic polyhedrosis virus segment 9 and its expression in BmN4 cells. *J Virol*, 1998, **72**( 7 ): 5762 ~ 5768.
- [ 8 ] Belloncik S, Liu J, Su D, et al. Identification and characterizations of a new cyovirus type 14, isolated from *Heliothis armigera*. *Journal of Invertebrate Pathology*, 1996, **67**: 41 ~ 47.

## Molecular Cloning and Sequencing of *Dendrolimus punctatus* Wenshanensis Cytoplasmic Polyhedrosis Virus ( DpwCPV ) NS5 Protein Gene

Wang Qiong Zhang Jiamin Wen Li Yang Juan Hu Yuanyang\*

( Institute of Virology, Wuhan University, Wuhan 430072, China )

**Abstract** : DpwCPV were purified from infected *Dendrolimus punctatus* Wenshanensis larvae by density gradient centrifugation. The genomic dsRNA segments were subsequently extracted directly from the purified polyhedra by SDS treatment, phenol/chloroform extraction and ethanol precipitation. Different segments of genomic dsRNA were separated by low melting point agarose electrophoresis, and S9 dsRNA was purified after being excised from the gel. S9 RNA was denaturalized in high temperature. The primers were designed on the basis of the RNA sequence of S9 in BmCPV-1 ( H strain ). After RT-PCR, the amplified cDNA was cloned into the pMD18-T Vector. The NS5 protein gene of DpwCPV is found to be 977 nucleotides in length with one open reading frame in one strand. It can code a predicted protein of 320 residues ( 35560 ).

**Key words** : *Dendrolimus punctatus* Wenshanensis T., *Dendrolimus punctatus* Wenshanensis Cytoplasmic polyhedrosis virus ( DpwCPV ), Protein NS5 gene

\* Corresponding author