

## 豆天蛾核型多角体病毒泛素基因的序列分析

何华纲, 朱姗姗, 王文兵\* (1. 江苏大学食品与生物工程学院, 江苏镇江 212013; 2. 江苏大学生命科学研究院, 江苏镇江 212013)

**摘要** 为了对 CbNPV ubiquitin 基因序列进行分析, 并对目前已知全基因组序列的杆状病毒的泛素氨基酸序列进行比较。从豆天蛾幼虫尸体中分离纯化豆天蛾核型多角体病毒(CbNPV), 提取基因组 DNA, 进行基因组测序。序列分析发现一个长度为 252 bp 的读码框序列与泛素基因同源性很高, 该基因编码 83 个氨基酸。氨基酸序列同源性分析结果表明, CbNPV 泛素的氨基酸序列与 20 种杆状病毒泛素的氨基酸序列的同源性在 66.3%~82.1%, 并且维持泛素功能所必需的氨基酸序列, 在大部分杆状病毒中也都保守存在。

**关键词** 豆天蛾核型多角体病毒; ubiquitin 基因; 序列分析

中图分类号 Q936 文献标识码 A 文章编号 0517-6611(2007)21-06376-03

Sequence Analysis of Ubiquitin Gene of *Clanis Bilineata* Nuclear Polyhedrosis Virus

HE Hua-gang et al (School of Food and Biological Engineering, Jiangsu University, Zhenjiang, Jiangsu 212013)

**Abstract** *Clanis bilineata* Nuclear Polyhedrosis Virus (CbNPV) particles were purified from *Clanis bilineata* body and then the genomic DNA was extracted. The viral DNA was sequenced with shotgun method. The result of sequence analysis indicated that the open read frame(ORF) of CbNPV ubiquitin contains 252 base pairs, which encoded 83 amino acids. Amino acid sequence analysis showed that CbUBI had 66.3%~82.1% identities with 20 baculovirus UBI proteins, and the residues known to be essential for the function of ubiquitin were conserved in most baculoviridae.

**Key words** CbNPV; Ubiquitin gene; Sequence analysis

泛素 Ubiquitin, Ubi) 是由 76 个氨基酸组成的, 具有高度保守性的小分子球状蛋白, 是所有真核细胞中普遍存在的最丰富的蛋白之一, 又称为遍在蛋白。泛素-蛋白酶体途径 (Ubiquitin-proteasome pathway, UPP) 是蛋白质选择性降解的主要方式, 它能催化泛素在泛素活化酶(E1)、泛素结合酶(E2)和泛素连接酶(E3)的作用下, 共价结合于靶蛋白上, 从而使经泛素化修饰的靶蛋白被 26S 蛋白酶体选择性降解成短肽<sup>[1]</sup>。除此之外, 泛素还参与了多种生命过程, 如 DNA 修复, 调节细胞周期进程, 转录和信号转导, 受体介导的内吞过程, 抗原提呈, 细胞的病理性改变及程序性死亡等<sup>[2]</sup>。同时, 近年研究还发现, 泛素在促进多种病毒增殖的过程中起着重要作用。

杆状病毒是有囊膜的双链 DNA 大形病毒, 宿主域仅限于无脊椎动物, 是可编码泛素或泛素样蛋白的病毒之一。笔者等从豆天蛾幼虫尸体中分离出豆天蛾核型多角体病毒 (*Clanis bilineata* Nuclear Polyhedrosis Virus, CbNPV), 并对其进行了全序列测定。在相似性比较分析中, 发现一长度为 252 bp 的读码框序列, 为 ubiquitin 类似基因。笔者即对 CbNPV ubiquitin 基因序列进行了分析, 并对目前已知全基因组序列的杆状病毒的泛素氨基酸序列进行了比较, 结果报道如下。

## 1 材料与方

1.1 材料 豆天蛾幼虫尸体采集于浙江省湖州市。

## 1.2 方法

1.2.1 CbNPV 多角体的纯化及 DNA 提取。感染病虫尸体研磨后过滤, 经蔗糖梯度离心纯化。将 CbNPV 悬于裂解液至清亮, 加入 SDS 和蛋白酶 K, 37℃ 下消化过夜。消化液分别经酚、酚氯仿、氯仿抽提后, 经乙醇沉淀得到 CbNPV 多角体

DNA。将沉淀溶于 2.0 mmol/L Tris (pH 值 8.0) 中, 4℃ 备用。

1.2.2 CbNPV 全基因组序列测定。采用 Shotgun 的方法构建 CbNPV 的 DNA 文库, 在上海人类基因组测序中心完成。

1.2.3 计算机分析 ubiquitin。测定的序列利用 BLAST 与 GenBank 数据库进行比较。用 DNASTar 软件包中的 MegAlign 软件对核苷酸及氨基酸序列进行同源性比较。

## 2 结果与分析

2.1 CbNPV ubiquitin 基因序列分析 在构建的 CbNPV DNA 文库中, 发现了 8 个包含 ubiquitin 基因序列的克隆(图 1)。对这些克隆进行多重测序分析, 获得了 ubiquitin 基因的完整序列(图 2)。

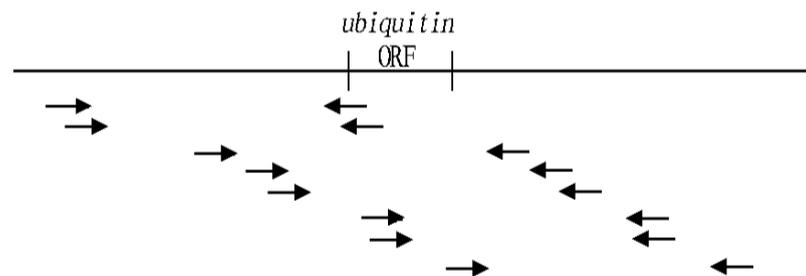


图1 与泛素基因相关的8个克隆片段

```
tatattgtacaattgtatgttaattgtgtacttagttataggacatgtaatttaataaatattttaa
tatttaataatttttaataacattctaaattattttatagataatttgtaacaaacggaaca
atgcaaatattttgtaagactttgaccgtaaatccatcacagtcgaagtgagtcgtct 60
M Q I F V K T L T G K S I T V E V E S S
gacactatcgagtcactgaagcaaaaaatcaactgacaagaaggattccacctgaccaa 120
D T I E S L K Q K I T D K E G I P P D Q
cagcggattattttgccggcaacaattagaagatgatcgactgtggcgattacaat 180
Q R I I F A G K Q L E D D R T V G D Y N
attcaaaaagaatccacgctacatttagtgtaaggctaaggaggaaatattataact 240
I Q K E S T L H L V L R L R G G N Y Y T
attaacaaataa 252
I N K
agtttatattttataaatatatttttttttttttttcaaacgggtatagataaat
```

注: 杆状病毒晚期基因转录的保守起始序列(ataag) 由方框标出。

维持泛素功能所必需的氨基酸残基以黑色字体标出。

## 图2 CbNPV ubiquitin 基因的核苷酸序列和推导的氨基酸序列

序列分析表明, CbNPV ubiquitin 基因由 252 个核苷酸组成, 编码 83 个氨基酸, 分子质量约为 9.5 kDa。转录起始位点上游 -27 bp 处具有晚期基因转录的保守起始序列 ataag, 表明该基因为晚期表达基因。

氨基酸序列分析表明, 真核生物中维持泛素功能所必需的氨基酸序列, 特别是 Lys<sup>48</sup> 和 His<sup>68</sup>, 以及 C 末端序列(LRL-

基金项目 江苏大学高级专业人才科研启动基金(05JDC033); 江苏省高校自然科学基金项目(06KJD180043); 江苏省研究生培养创新工程项目。

作者简介 何华纲(1978-), 男, 江苏武进人, 助教, 从事病毒分子生物学研究。\* 通讯作者, E-mail: wenbingwang@ujs.edu.cn。

收稿日期 2007-04-03

RGG) 在 CbUBI 中均保守存在。另外, 在 CbNPV UBI 保守序列 LRLRGG 后多出由 7 个氨基酸组成的延伸肽。

**2.2 ubiquitin 基因氨基酸序列同源性比较** 应用 DNASar 软件包中的 MegAlign 软件对 CbNPV 与人、酵母、家蚕及 20 种

杆状病毒泛素的氨基酸序列进行同源性比较。结果表明, CbNPV UBI 与人、酵母和家蚕泛素的氨基酸序列的同源性均为 85.5%, 与 20 种杆状病毒泛素的氨基酸序列的同源性在 66.3% ~ 82.1% (图 3)。

		Percent Identity																								
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
Divergence	1	■	96.1	98.7	76.3	86.8	76.3	85.5	78.9	80.3	76.3	76.3	77.6	82.9	80.3	76.3	84.2	84.2	80.3	81.6	81.6	77.6	78.9	80.3	Human	
	2	4.1	■	96.1	75.0	86.8	75.0	85.5	77.6	80.3	75.0	76.3	76.3	77.6	80.3	78.9	75.0	84.2	81.6	78.9	80.3	81.6	77.6	80.3	81.6	Yeast
	3	1.3	4.1	■	75.0	86.8	75.0	85.5	77.6	80.3	75.0	76.3	76.3	78.9	81.6	78.9	75.0	84.2	82.9	78.9	80.3	81.6	78.9	78.9	80.3	Bm
	4	28.5	30.4	30.4	■	68.8	100.0	74.0	87.0	67.5	86.8	74.0	74.0	70.1	77.9	84.4	100.0	74.0	79.2	67.5	78.9	76.6	68.8	76.6	77.9	AcNPV
	5	14.5	14.5	14.5	38.7	■	68.8	75.0	74.4	75.6	72.4	72.5	72.5	73.8	73.8	73.8	68.8	77.5	77.5	73.8	78.9	72.5	71.2	73.8	76.6	AhNPV
	6	28.5	30.4	30.4	0.0	38.7	■	74.0	87.0	67.5	86.8	74.0	74.0	70.1	77.9	84.4	100.0	74.0	79.2	67.5	78.9	76.6	68.8	76.6	77.9	BmNPV
	7	16.1	16.1	16.1	30.4	24.8	30.4	■	74.4	82.1	75.0	68.7	68.7	69.9	75.9	72.3	74.0	80.0	74.7	66.3	77.6	73.5	66.3	74.7	80.5	CbNPV
	8	24.8	26.6	26.6	12.9	28.5	12.9	28.5	■	74.4	93.4	78.2	78.2	73.1	80.8	94.9	87.0	80.8	80.8	69.2	80.3	79.5	74.4	74.4	77.9	CfNPV
	9	23.0	23.0	23.0	40.9	26.6	40.9	17.8	30.4	■	76.3	67.9	67.9	67.9	79.5	73.1	67.5	83.3	71.8	71.8	72.4	73.1	67.9	71.8	70.1	ChchNPV
	10	28.5	30.4	30.4	14.5	34.5	14.5	30.4	6.9	28.5	■	76.3	76.3	71.1	78.9	90.8	86.8	78.9	80.3	69.7	77.6	78.9	72.4	73.7	76.3	EpNPV
	11	28.5	28.5	28.5	30.4	28.5	30.4	30.4	23.0	38.7	28.5	■	98.8	75.9	71.1	74.7	74.0	75.0	81.9	63.9	77.6	72.3	68.7	71.1	79.2	HaNPV
	12	28.5	28.5	28.5	30.4	28.5	30.4	30.4	23.0	38.7	28.5	0.0	■	77.1	71.1	74.7	74.0	75.0	80.7	63.9	77.6	72.3	68.7	71.1	79.2	HzNPV
	13	26.6	26.6	24.8	36.6	26.6	36.6	28.5	30.4	38.7	36.6	19.5	19.5	■	62.0	63.4	70.1	75.0	42.7	60.6	82.9	66.0	66.0	67.7	84.4	LdNPV
	14	19.5	23.0	21.2	24.8	26.6	24.8	19.5	19.5	21.2	24.8	26.6	26.6	24.8	■	67.7	77.9	90.0	67.0	58.5	77.6	63.8	61.7	62.4	76.6	MacoNPV
	15	23.0	24.8	24.8	16.1	26.6	16.1	28.5	2.7	30.4	9.8	24.8	24.8	28.5	21.2	■	84.4	77.5	68.8	59.1	81.6	67.7	63.4	64.5	79.2	OpNPV
	16	28.5	30.4	30.4	0.0	38.7	0.0	30.4	12.9	40.9	14.5	30.4	30.4	36.6	24.8	16.1	■	74.0	79.2	67.5	78.9	76.6	68.8	76.6	77.9	RoNPV
	17	17.8	17.8	17.8	30.4	23.0	30.4	17.8	19.5	16.1	24.8	24.8	24.8	24.8	5.5	21.2	30.4	■	80.0	70.0	78.9	77.5	75.0	73.8	77.9	SeNPV
	18	17.8	21.2	19.5	23.0	24.8	23.0	23.0	19.5	32.4	23.0	12.9	12.9	24.8	17.8	21.2	23.0	19.5	■	62.8	82.9	68.1	63.8	67.7	81.8	SpltNPV
	19	23.0	24.8	24.8	40.9	26.6	40.9	34.5	36.6	32.4	38.7	38.7	38.7	30.4	34.5	34.5	40.9	34.5	26.6	■	80.3	73.4	75.5	65.6	75.3	AoGV
	20	21.2	23.0	23.0	24.8	24.8	24.8	26.6	23.0	34.5	26.6	26.6	26.6	19.5	26.6	21.2	24.8	24.8	19.5	23.0	■	89.5	84.2	85.5	88.2	CIGV
	21	21.2	21.2	21.2	26.6	28.5	26.6	23.0	21.2	30.4	24.8	24.8	24.8	21.2	26.6	19.5	26.6	23.0	17.8	30.4	11.4	■	80.9	76.3	83.1	CpGV
	22	26.6	26.6	24.8	38.7	30.4	38.7	34.5	28.5	38.7	34.5	30.4	30.4	21.2	30.4	26.6	38.7	26.6	24.8	26.6	17.8	23.0	■	67.7	81.8	PoGV
	23	24.8	23.0	24.8	26.6	26.6	26.6	23.0	28.5	32.4	32.4	26.6	26.6	21.2	28.5	26.6	26.6	26.6	21.2	28.5	16.1	12.9	26.6	■	87.0	PxGV
	24	23.0	21.2	23.0	24.8	26.6	24.8	21.2	24.8	36.6	28.5	24.8	24.8	16.1	26.6	23.0	24.8	24.8	19.5	28.5	12.9	17.8	19.5	12.9	■	XcGV
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24		

图3 人、酵母、家蚕和 21 种杆状病毒泛素基因氨基酸序列的同源性比较

### 3 讨论

泛素的氨基酸序列在真核生物中高度保守, 动物、酵母和植物中仅有 3 个氨基酸的差别。泛素编码序列可分为 2 种类型: 多聚泛素基因编码由多个首尾相接的泛素序列组成的多聚泛素, 多聚泛素在去泛素酶(E4) 的作用下可游离出多个泛素单体。另一类 ubiquitin 基因编码的泛素序列, 在其 C 端连接有核糖体蛋白, 或称为 C 端延伸肽(Carboxyl extension protein, CEP) 序列<sup>[3]</sup>。CEP 也可以在去泛素酶(E4) 的催化下从融合蛋白上释放。已有研究表明, 泛素在多种病毒的发育循环中起着重要的作用。如宿主的泛素分子可共价连接在一些植物病毒的外壳蛋白上<sup>[4]</sup>, 游离的泛素也可作为禽白血病毒病的组成成分大量存在<sup>[5]</sup>, 非洲猪瘟病毒(ASFV) 可编码泛素连接酶(E3) 样蛋白等<sup>[6]</sup>。但目前所知能够编码泛素或泛素样蛋白的病毒只有杆状病毒和昆虫痘病毒。

杆状病毒是一类主要以鳞翅目昆虫为宿主的, 有囊膜的双链 DNA 病毒, 是农林重要害虫的自然控制因子<sup>[7]</sup>。苜蓿尺蠖核型多角体病毒(*Autographa californica Nuclear Polyhedrosis Virus*, AcNPV) 是第一个被报道可编码泛素样蛋白的杆状病毒, 和动物细胞的泛素蛋白相差 18 个氨基酸。它是 AcNPV 出芽病毒粒子的结构蛋白, 通过与磷脂的共价连接而锚定在病毒粒子囊膜内侧<sup>[8]</sup>。目前已公布全基因组序列的 24 种杆状病毒中, 除了黄地老虎颗粒体病毒(*Agrotis segetum granulovirus*, AgsGV), 红头松树叶蜂核型多角体病毒(*Neodiprion lecontei NPV*, NleNPV), 松柏锯角叶蜂核型多角体病毒(*Neodiprion sertifer NPV*, NeseNPV) 和库蚊核型多角体病毒(*Culex nigripalpus NPV*, CuriNPV) 外, 其余的杆状病毒都具有可以编码泛素样蛋白的序列。在真核生物中维持泛素功能所必需的氨基酸序列, 大部分杆状病毒中也都保守存在。

如, 参与形成泛素-蛋白酶复合体的 4 个氨基酸残基 Lys<sup>29</sup>、Lys<sup>48</sup>、Lys<sup>63</sup>、His<sup>68</sup>, 除了 Lys<sup>29</sup> 外, 其余在杆状病毒 UBI 中均保守存在。Lys<sup>29</sup> 在苹浅褐卷蛾核型多角体病毒(*Epiphyas postvittana NPV*, EpNPV)、斜纹夜蛾核型多角体病毒(*Spodoptera litura NPV*, SpltNPV)、棉褐带卷蛾颗粒体病毒(*Adoxophyes orana GV*, AoGV) 和马铃薯块茎蛾颗粒体病毒(*Plutella maculipennis operculella GV*, PoGV) 中分别替换为 Gln、Pro、Ala 和 Leu; 在受体介导的内吞作用中起重要作用的 Gln<sup>2</sup>、Phe<sup>4</sup>、Lys<sup>6</sup>、Leu<sup>8</sup>、Thr<sup>12</sup>、Ile<sup>44</sup>、Glu<sup>64</sup> 及 Val<sup>70</sup><sup>[9]</sup>, 除了 Thr<sup>12</sup> 在 CbNPV、戈尔登双生斑核型多角体病毒(*Chrysodixis chalcites NPV*, ChchNPV) 和 EpNPV 中替换为 Ser, 以及 Glu<sup>64</sup> 及在茶小卷叶蛾核型多角体病毒(*Adoxophyes honmai NPV*, AhNPV) 中替换为 Asp 外, 其余均保守存在; 对于泛素分子维持正确构象所必需的 C 末端保守序列 LRLRGG 则在所有已知全基因组序列的杆状病毒中都保守存在。

氨基酸序列同源性分析表明(图 4): 杆状病毒泛素的氨基酸序列同源性在 58.5% ~ 100%, 变化范围较大, 这主要是由于杆状病毒编码的泛素蛋白, 在其 C 端均包含有 1 ~ 256 个氨基酸组成的长短不等的 CEP, CEP 的氨基酸序列在各杆状病毒间并不保守。推测它们可能来源于祖先细胞的泛素融合基因或多聚泛素基因的 C 末端延长部分<sup>[10]</sup>。另外, 杆状病毒泛素的氨基酸序列中存在 2 个异质性相对较高的区域: HR1(15 ~ 31 位) 和 HR2(53 ~ 57 位), 可能与宿主特异性有关(图 3)。杆状病毒泛素间的氨基酸序列保守程度较真核的泛素要低很多, 说明前者在功能上可能与后者存在差异。移码突变实验表明, AcNPV 编码的 UBI 蛋白对于病毒的复制过程是一种非必需蛋白, 但泛素作为杆状病毒出芽病毒粒子的结构蛋白之一, 共价结合于囊膜内侧, 推测泛素与病毒颗粒的装配或出芽过程有关, UBI 锚定于膜上也可能是病毒感染

