

SINPV PK 基因的部分序列分析*

魏永杰 龙蔡新 陈尚武 王珣章

(中山大学生物防治国家重点实验室, 广州 510275)

摘要 *SINPV* 基因组 DNA 的 *Hind*III-*Xba*I 片段上含有部分的 PK 基因. *SINPV* PK 的氨基酸序列与 *HaNPV*, *HzNPV*, *SpliNPV*, *AfNPV*, *AcNPV*, *LdNPV* PK 氨基酸序列的同源性分别为 45%, 44%, 89%, 37%, 38% 和 37%, 其上含有蛋白激酶的特征序列 IVHANDVKLEN-VL.

关键词 *SINPV*, PK 基因, 共有序列

分类号 Q 751

蛋白质的可逆磷酸化可调节蛋白质在细胞内的分布及胞内酶的活性, 细胞代谢、细胞周期、离子运输等细胞活动都有蛋白质磷酸化过程的参与. 蛋白质的这种可逆磷酸化过程由蛋白激酶 (protein kinase, PK) 及磷酸酯酶 (phosphatase) 的相互作用来完成. 杆状病毒中, 蛋白激酶 (PK) 的活动与多角体的形成有关, 且对病毒 DNA 从病毒粒子中的释放至关重要^[1]. 杆状病毒的 PK 基因属滞早期基因, 在病毒生活周期的晚期和极晚期也有持续表达^[2]. 斜纹夜蛾核多角体病毒 (*Spodoptera litura* nuclear polyhedrosis virus, *SINPV*) 能感染并致死斜纹夜蛾 (*Spodoptera litura*), 对斜纹夜蛾具有较好的防治效果. 本研究报道了 *SINPV* PK 基因的部分序列, 并对这一序列进行了计算机分析.

1 材料与方法

(1) DNA 片段的序列测定: 参照美国 BIOCHEMICAL 公司 SEQUENCE KIT 操作手册, 测定重组质粒 pP18PH3 上插入片段的序列, 重组质粒 pP18PH 的构建过程详见另文所述 (魏永杰等, 待发表).

(2) DNA 序列分析: 分析软件为 GCG package (V 8.0), 所用数据库为 GenBank Prosite Transfac 数据库.

2 结果分析

核多角体病毒 (nuclear polyhedrosis virus, NPV) 的多角体基因及 PK 基因在位置上具有相关性, GenBank/EMBL 数据库中所有 NPV 的 PK 基因均处在各自多角体基因下游 1 250~1 700 bp 处. 据此, 本文推测 *SINPV* 的 PK 基因位于多角体基因下游的基因组 DNA 1 200 bp 的 *Hind*III-*Xba*I 片段上, 该片段已被克隆到载体质粒 pUC18 上构建成重组质粒

* 国家自然科学基金 (39730030) 资助项目

文中 PK 基因序列已被 GenBank 收录, 收录号为 AF 039272

收稿日期: 1998-01-15 魏永杰, 男, 25岁, 博士研究生

pP18PH3 (魏永杰等, 待发表).

SINPV 基因组 DNA 1 200 bp 的 *Hind*III-*Xba*I 片段的序列分析结果如图 所示. 40 位的 ATG 是 PK 基因的起始密码子, ATG 周围序列为 ATCATGG, 符合典型的 Kozak 规则. 起始密码子上游 15 bp 处, 在 386 位的 AGCGTT 为杆状病毒早期基因的 CGT 基序, 其共有序列为 AXCGT (G/T)^[31]. 35 位及 36 位的互补链上各有 1 个 TATA 盒; 168 位处的 CG-GATAGA 是与昆虫细胞中 GATA 样蛋白结合的 GATA 基序, 这类基序在杆状病毒其它一些早期基因, 如 *gp64* 基因中也有存在^[31]. PK 基因的上游序列中未发现晚期基因的 TAAG 基序, 这与 *LdNPV* PK 基因有所不同^[21].

```

1  A  GCT TCA AAA GTG ACA AGA CAC GTT CCA AAT GTT TGT TGA CTT GTA CCG GAT CGT 55
56 GGG ATT TCT TTG GGA GAT CAA TGT GTT TTT TTC ATG GTA AGT TCG TAT TGT CTG TCA 112
113 CAT TGC ACG GCA TCG ATG GCT CTA TCG ATC TGG ATA ATT TTG TTT AGT AAA GAT 166
167 TCGGATAGA GCC GGA CGT ACT GAA ACT TTA TCT TCG TAC ATG TCA TCG GCT ATT CGC 223
224 AGC AGT TGC ATG CAT TCG TAA ATA TCT AAA TTC ACC ACT TTT GAT TGC GGA CTC ACC 280
281 GAG ATC AGT TTG CGT TTC AGT TCG ACA CAG GGC GAC GGT GTC ACG AGT CGC GAC ATT 337
338 AAC AAT TTT AAA TCT ACATTA TATA TTA TTC GCT TTG AAA AGT TCA GCC AGCGTT GCC 394
397 GCC ATC ATG GAC GCC CGC TTT GTT GAA GAT ATC AAT CAG TTT TGT AGA GAC GTA CAA 451
1  M  D  A  R  G  V  E  D  I  N  Q  F  C  R  D  V  Q  17
452 GTT CAC TCT TCC TGC CGT TTG ATA GAC GGC AAG TAC GGC AAA GTG AGC GTG ATT AGA 508
18  V  H  S  S  C  R  L  I  D  G  K  Y  G  K  V  S  V  I  R  36
509 CAC ACT CCG ACT GAC AAG TTG TTT TTG AAA AAG ATA ATC AAG CAC GAA AAC TTT GAT 565
37  H  T  P  T  D  K  L  F  L  K  K  I  I  K  H  E  N  F  D  55
566 GCG ATC GAA TTC ACG GTG CAT CAC CTT ATG AGA GAC CAT TCG AAT TTC GTT AGA CTC 622
56  A  I  E  L  T  G  V  H  H  L  M  R  D  H  S  N  F  V  R  L  74
623 TAT TAT GCG TTT AAC ACG CCA AAC ATA AAC GTG CTG ATC ATG GAC TAT GTC GAG GAC 679
75  Y  Y  A  F  N  T  P  K  I  N  V  L  I  M  D  Y  V  E  D  93
680 GGA GAT CTG TTT GTT ATA TTA AAG ACT GAC GGG CCG TTT AGT GAA CGC AAG ACT GCT 736
94  G  D  L  F  E  I  L  K  T  D  G  P  F  S  E  R  K  T  A  112
737 CAT GTA GTG TCT CAA TTG GCC TTT GCT CTT AAC TCT TTA CAC GGA CAC AGG ATCGTT 793
113 H  V  V  S  Q  L  A  F  A  L  N  S  L  H  G  H  R  I  V  131
794 CAC AAC GAT GTA AAA TTG GAA AAC GTT TTG TTT ACT AAG AGC AAT AGA CAA GTT AAA 850
132 H  N  D  V  K  L  E  N  V  L  F  T  K  S  N  R  Q  V  R  K  150
851 CTG TGC GAT TAC GGT CTA TGC AAA CCG ATC GGG GCC GTA CAG AAA CAA GAC GGC ACC 987
151 L  C  D  Y  G  L  C  K  P  I  G  A  V  Q  K  Q  D  G  T  169
908 GTG GAT TAC TTT TCT CCT GAA AAG ATT CGA GAC AAG CCG TAC GAC GAG ACC ATG GAC 964
170 V  D  Y  F  S  P  E  K  I  R  D  K  P  Y  D  E  T  M  D  188
965 TTT TGG GCC ATC GGC GTG ATT GCA TAC GAA CTG TTG ACT GGC GAG CAT CCA TAC AAG 1021
189 F  W  A  I  G  V  I  A  Y  E  L  L  T  G  E  H  P  Y  K  207
1022 GAT GGC AAA GAT GAC GAC TTG GAC GTC GAC ACT TTA AAG TAT CGT ATT GAAAAA AGT 1078
208 D  G  K  D  D  D  L  D  V  D  T  L  K  Y  R  I  E  K  S  226
1079 TTT AGA CCT ATA CAA AAG CGT AAC ATA TCC GAG ACC GCT TTG GAC TTT CTC AAG AAC 1135
227 F  R  P  I  Q  K  R  N  I  S  E  T  A  L  D  F  L  K  N  245
1136 ATG TTG CAT CCC AAC ATT AAA TCT AGA 1162
246 M  L  H  P  N  I  K  S  R  254

```

图 1 *SINPV* 的 1 200 bp *Hind*III-*Xba*I 片段核苷酸序列及其编码的部分蛋白激酶的氨基酸序列

Fig. 1 Nucleotide sequence of the 1 200 bp *Hind*III-*Xba*I fragment and the predicted amino acid sequence of partial protein kinase

框线: 早期基因 CGT 基序, GATA 基序及蛋白激酶特征氨基酸序列; 单线: TATA box

SINPV 蛋白激酶的氨基酸序列与 *Helioverpa armigera* (*Ha*) NPV, *Helioverpa zea* (*Hs*) NPV, *Spodoptera littoralis* (*Spli*) NPV, *Anagrapha falcifera* (*Af*) NPV, *Autographa californica* (*Ac*) NPV, *Lymantria dispar* (*Ld*) NPV 蛋白激酶氨基酸序列的同源性分别为 45%, 44%, 89%, 37%, 38% 和 37%。其中 *SINPV* 与 *Spli* NPV 蛋白激酶的氨基酸序列

同源性最高, 达 89%, 进一步证明这两种病毒有较近的亲缘关系^[4].

SINPV 蛋白激酶 N 端 130~142 位的 IV HAN DV K L E N V L 为蛋白激酶的特征序列^[5], 这段序列在各 NPV 的蛋白激酶中保守. 另外, N 端 152~158 位的 CDYGLCK 也有很高的保守性, 表明这段序列可能是 NPV 类蛋白激酶的重要功能区.

PK 基因为杆状病毒的必需基因; 蛋白激酶 (PK) 参与病毒多种结构蛋白与调节蛋白的磷酸化, 与多角体的形成及病毒粒子的释放密切相关. 对 PK 基因的研究有助于了解 *SINPV* 的侵染、复制、基因的表达与调控及 *SINPV* 与宿主细胞的相互作用, 并为利用 *SINPV* 防治斜纹夜蛾提供理论基础.

参 考 文 献

- 1 Miller L K, Adang M J, Browne D. Protein kinase activity associated with the extracellular and occluded forms of the baculovirus *Autographa californica* nuclear polyhedrosis virus. *J Virol*, 1983, 46: 275~278
- 2 Bischoff D S, Slavicek J M. Identification and characterization of a protein kinase gene in the *Lymantria dispar* multinucleocapsid nuclear polyhedrosis virus. *J Virol*, 1994, 68: 1728~1736
- 3 Blissard G W, Rohrmann G F. Location, sequence, transcriptional mapping and temporal expression of the *gp64* envelope glycoprotein gene of *Orgyia pseudotsugata* multicapsid nuclear polyhedrosis virus. *Virology*, 1989, 176: 537~555
- 4 Maeda S, Mukohara Y, Kondo A. Characteristically distinct isolates of the nuclear polyhedrosis virus from *Spodoptera litura*. *J Gen Virol*, 1990, 71: 2631~2639
- 5 Hanks S K, Quinn A M, Hunter T. The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science*, 1988, 241: 42~52

Partial Nucleotide Sequence of Protein Kinase Gene of *Spodoptera litura* Nuclear Polyhedrosis Virus

Wei Yongjie* Long Qixin Chen Shangwu Wang Xunzhang

Abstract The partial protein kinase gene of *Spodoptera litura* nuclear polyhedrosis virus was sequenced. The predicted *SINPV* protein kinase shows 43%, 44%, 89%, 37%, 38% and 37% amino acid homology to those of *HaNPV*, *HZNPV*, *SpliNPV*, *AfNPV*, *AcMN-PV* and *LdNPV*, respectively. There is an IV HAN DV K L E N V L pattern in the protein kinase of *SINPV*, which is the signature motif of protein kinase family.

Keywords *Spodoptera litura* (*Sl*) nuclear polyhedrosis virus, protein kinase, consensus sequence

* State Key Laboratory for Biocontrol, Zhongshan University, Guangzhou 510275, China