

苏云金杆菌 *Kurstaki* HD-1 亚种的一个杀虫蛋白 基因毒性区的核苷酸序列*

乔利亚 田颖川 莽克强

(中国科学院微生物研究所,北京 100080)

为了更好地利用苏云金芽孢杆菌(*Bacillus thuringiensis*,以下简称 *B. t*)这个生物杀虫剂并培育抗虫转基因植物,80年代以来各国科学家们纷纷加强了对 *B. t* 杀虫蛋白基因的结构和功能方面的研究。到目前为止已有 42 种以上的 *B. t* 杀虫蛋白基因(或称为晶体蛋白基因)得到克隆并进行了 DNA 序列分析。我们已克隆了 *B. t kurstaki* HD-1 亚种的晶体蛋白(Cry)基因,经酶切图谱分析将其中两个克隆的 Cry 基因,按 Kronstad 和 Whiteley^[1] 的分类方法分别确定为 5.3kb 和 6.6kb 类型的 Cry 基因^[2,3]。按 Hofte 和 Whiteley^[4] 1989 年的新分类系统,这两个基因应分别属于 Cry I A(b)和 Cry I A(c)两个基因型。本文首次报道 *B. t* HD-1 的 Cry I A(c)基因编码杀虫蛋白毒性区 2.1kb 的核苷酸序列并与已发表的 *B. t* HD-73 的 Cry I A(c)^[5] 及 *B. t* HD-1 的 Cry I A(b)^[6] 基因的序列进行比较。

材料与 方法

(一) 质粒

1. 质粒 pB48.102 为 pB48.101^[3]经 Kpn I 酶解后去掉此 *B. t* Cry I A(c)基因 3'端 Kpn I 下游片段的质粒。pB48.102 含有 *B. t* Cry I A(c)基因 5'端 2.1kb 的序列。
2. 质粒 pBluescript KS⁻ 或 KS⁺ 系美国 Stratagene 公司产品。

(二) 生化试剂

1. 限制性内切酶及核酸修饰酶购自 Promega、BioLabs 或 Boehringer 公司。
2. DNA 序列测定盒系 USB 及 Pharmacia 公司产品。
3. α -³²P-dCTP 和 α -³⁵S-dATP 为美国 NEN 公司产品。

(三) DNA 序列分析

以 pB48.102 为起始质粒利用 Cry I A(c)基因内的限制性内切酶 EcoR I、Xba I、EcoRV、Cla I、Pst I 等酶切位点将此 2.1kb 基因构建成互相重叠的 17 个亚克隆。运用 Sanger 的双脱氧法按试剂盒厂家提供的方法步骤对亚克隆进行序列分析。

结果和 讨论

(一) *Bacillus thuringiensis* (*B. t*) Cry I A(c)-48 基因毒性区的 DNA 序列

B. t HD-73 的杀虫蛋白基因属 Cry I A(c)基因型^[4],为了比较方便,我们将本实验室克隆、测序的 *B. t* HD-1 Cry I A(c)基因定名为 Cry I A(c)-48。将 *B. t* HD-73 的杀虫蛋白基因定名为 Cry I A(c)-73。我们测定了 Cry I A(c)-48 基因编码区 1-2169bp 的核苷酸序列,结果表明其序列与 Cry I A(c)-73 的基

本文于 1992 年 7 月 2 日收到。

* 本工作系国家“七五”攻关项目并得到国际科学文化中心世界实验室(WL, ICSC,日内瓦,洛桑)的部分资助。感谢方荣祥先生在测序工作中的指导和帮助。

表1 苏云金杆菌杀虫蛋白基因Cry I A(c)-48毒性区的密码子使用情况

密码子数	密码子	氨基酸	密码子百分数	密码子数	密码子	氨基酸	密码子百分数	密码子数	密码子	氨基酸	密码子百分数	密码子数	密码子	氨基酸	密码子百分数
32	UUU	Phe	4.4	14	UCU	Ser	1.9	26	UAU	Try	3.6	2	UGU	Cys	0.3
7	UUC	Phe	1.0	9	UCC	Ser	1.2	5	UAC	Try	0.7	1	UGC	Cys	0.1
24	UUA	Leu	3.3	14	UCA	Ser	1.9	0	UAA	End	0.0	0	UGA	End	0.0
7	UUG	Leu	1.0	7	UCG	Ser	1.0	0	UAG	End	0.0	11	UGG	Trp	1.5
10	CUU	Leu	1.4	10	CCU	Pro	1.4	9	CAU	His	1.2	8	CGU	Arg	1.1
4	CUC	Leu	0.6	2	CCC	Pro	0.3	2	CAC	His	0.3	2	CGC	Arg	0.3
11	CUA	Leu	1.5	17	CCA	Pro	2.3	27	CAA	Gln	3.7	6	CGA	Arg	0.8
5	CUG	Leu	0.7	5	CCG	Pro	0.7	6	CAG	Gln	0.8	1	CGG	Arg	0.1
26	AUU	Ile	3.6	10	ACU	Thr	1.4	41	AAU	Asn	5.7	23	AGU	Ser	3.2
7	AUC	Ile	1.0	8	ACC	Thr	1.1	15	AAC	Asn	2.1	3	AGC	Ser	0.4
18	AUA	Ile	2.5	17	ACA	Thr	2.3	6	AAA	Iys	0.8	24	AGA	Arg	3.3
8	AUG	Met	1.1	9	ACG	Thr	1.2	4	AAG	Iys	0.6	8	AGG	Arg	1.1
19	GUU	Val	2.6	20	GCU	Ala	2.8	27	GAU	Asp	2.1	15	GGU	Gly	2.1
3	GUC	Val	0.4	4	GCC	Ala	0.6	7	GAC	Asp	1.0	5	GGC	Gly	0.7
19	GUA	Val	2.6	11	GCA	Ala	1.5	33	GAA	Glu	4.6	22	GGA	Gly	3.0
8	GUG	Val	1.1	5	GCG	Ala	0.7	4	GAG	Glu	0.6	10	GGG	Gly	1.4

	1		442		976		1000
A:	ATG	TTG	GTC GGT	ACG ...			
		Leu	Val Gly	Thr			
B:	ATG	TTT	GTA GGG	ACT ...			
		Phe	Val Gly	Thr			
	1018		1087		1240		
A:	ACC	TCC TCT ACT TTT	CCA				
	Thr	Ser Ser Phe	Pro				
B:	ACT	TCG TCC ACT TTA	CCG				
	Thr	Ser Ser Leu	Pro				
	1318		1777		2101		
A:	TCT AGT AGT AGT GTA	GGA	GGT ...				
	Ser Ser Ser Ser Val	Gly	Gly				
B:	TTT AGT AAT AGT AGT GTA	GGG	GGC ...				
	Phe Ser Asn Ser Ser Val	Gly	Gly				

图1 Cry I A(c)-48 基因与 Cry I A(c)-73 基因毒性区序列的比较

A: Cry I A(c)-48 B. t 毒蛋白基因的部分序列; B: Cry I A(c)-73 的部分序列;
虚线表示两个基因的核苷酸序列完全相同的部分

MDNNPNINECIPYNCLSNPEVEVLGGERIEIGYIPIDISLSLTQFLSEFVPGAGFVGLVDIIW *****	65
GIFGPSQWDAFLVQIEQLINQRIIEEFARNQATSRLEGLSNLYQIYAESFREWEADPTNPALREEM *****	130
RIQFNDMNSALITAIPLLAVQNYQVPLLSVYVQAANLHLSVLRDVSYFGQRWGFDAATINSRYND *****F*****	195
LTRLIGNYTDYAVRWYNTGLERVWGPDSRDVWRVYNQFRRELTLYLDIIVALFPNYDSRRYPJRTY *****H*****I*****S*****J*****	260
SQIPREIYTNPVLENFDGSAOGIERSIRSPHLMNDILNSITTYDAHRGYYYWSGHQINASP *****G*****E*****	325
YVGFSGPEFTFPLYGTNGNAAPQORIVAQLGQGVYRTLSSTFYRRPFNIGINQQLSVLDGTEFAY *****L*****	390
GTSSNLPASVYRKSGTYDSDLDEIPPNONNVPPROGFSNRLSHVSMFRSGSS SSVSIIRAPNFSWI *****F*N*****	455
HRSAEFNIIASDSITQIPAVKGNFLFNGSVISGPGFTGGDLVRLNSSGNNIQNRGYIEVPIHFS *****P*SQ*****LT*STN*GS*TSVVKGPQFT*GDILRRT*PCQ*STLRVNIITAPLSQ	520
STSTRYRVRYVRYASVYPIHRNVNWNSSIFSNYVATATSCDNLQSSDFCYFESANAFTSSLGNI RYRY*1*YASTTNLQFHITSIDGRPI*QGN**A*NSSGSNLQSGSFRITVGFITPFNFNSNG**VFTL	585
VGVRNFSGTAGVVIDRFEFIPVTAILEAEYNLERAQKAVNALFTSTNQLGLKTNVTDYHIDQVSN SAHYVFN**NEBY* **I**V*ATV*F****D****R***E*****I****D*****	650
LVTYLSDEFCLDEKRELSKVKHAKRLSDERNLLODSNFKDINRQPERGWGSGTIGITIQGGDDVF **EC*****K*****P**RG***LD***R***D*****	715
KENYVTLS *****	723

图2 Cry I A(c)-48(上行)与Cry I A(b)(下行)基因毒性区编码的氨基酸序列比较

* 表示与上行相同的氨基酸

本相同(资料略)。对几个 *B. t* Cry 基因缺失分析的结果表明, *B. t* 杀虫蛋白的毒性是由前毒素蛋白(130kd 左右)N 端一半决定的^[4], 最短的毒性片段为该蛋白的第 29 到第 607 个氨基酸的部分。 *B. t* HD-73 杀虫蛋白基因[Cry I A(c)-73]^[5]在 3' 缺失到 1836bp 处仍能在大肠杆菌中表达出有杀虫活性的杀虫蛋白。这里所报道的序列编码杀虫蛋白 N 端 723 个氨基酸, 完全包括了该基因所编码的毒性区。与其它原核生物的基因类似, Cry I A(c)-48 也富含 AT, 在这段毒性区内 AT 占 62%。表 1 列出这段基因的氨基酸密码子使用情况, 从三联密码子的兼并性看, 75% 的密码子的第三位上选择了 A 或 T, 这与 Cry I A(c)-73 基因^[5]的情况一致。

(二) Cry I A(c)-48 基因与 Cry I A(c)-73 及 Cry I A(b)^[6]基因的比较

杀虫蛋白基因 Cry I A(c)-48 与 Cry I A(c)-73 的核苷酸序列比较结果(见图 1)表明这两个属同一基因型的基因编码毒性蛋白区的序列几乎是完全一样的, 同源性达 99%, 在 2169 个碱基对中只有 12 个

不同,包括 Cry I A(c)-48 在 1323bp 后有三个碱基缺失(AAT),共有 15 个碱基的差异,这些差异引起氨基酸的不同有三处,即第 148 位(442-444bp),在 Cry I A(c)-48 中的为 Leu,而在 Cry I A(c)-73 中为 Phe,第 366 位(1096-1098bp),Cry I A(c)-48 中为 Phe,而在 Cry I A(c)-73 中为 Leu,以及第 440 位(1318-1320bp),Cry I A(c)-48 中为 Ser,而在 Cry I A(c)-73 中为 Phe。与 Cry I A(c)-73 相比,Cry I A(c)-48 在第 441 个氨基酸后因碱基缺失而少掉一个氨基酸(Asn)。这四个不同的氨基酸主要位于此毒性区中部靠近 N 端一半,在 C 端两者所编码的氨基酸序列完全一致。根据毒性蛋白存在两个区(即毒性区和特异性区)的设想^[7]毒蛋白的 N 端一半为毒性区,与杀虫毒力有关;而 C 端一半为特异性区,与毒蛋白与昆虫肠膜的受体位点特异结合有关。所以推测 Cry I A(c)-48 与 Cry I A(c)-73 的杀虫蛋白在毒力方面可能会略有差异,而寄主范围应是一样的。

Cry I A(c)-48 基因与 Cry I A(b)基因^[6]编码的氨基酸序列的比较结果见图 2。这两个基因的杀虫蛋白在 N 端(第 1-455 氨基酸)有 10 个氨基酸差异,所以它们在杀虫毒性方面可能略有不同。而 C 端(氨基酸 456-723)有 131 个氨基酸差异,这种差异可能造成这两种蛋白在昆虫膜上对受体的竞争力上有所不同。

Cry I A(c)-48 基因毒性区 DNA 序列的完成为进一步改造和利用此杀虫蛋白基因打下了基础。

参 考 文 献

- [1] Krenstad, J. W. and H. R. Whiteley; *Gene*, **43**:29-40,1986.
- [2] 田颖川等:生物工程学报 **5**, (1):1-10,1989.
- [3] 田颖川等:生物工程学报 **7**, (1):1-10,1991.
- [4] Hofte, H. and H. R. Whiteley; *Microbiological Reviews*, **53**(2): 242-255, 1989.
- [5] Adang, M. J. et al.; *Gene*, **36**:289-300,1985.
- [6] Geiser, M. et al.; *ibid.*, **48**:109-118,1986.
- [7] Choma, C. T. et al.; *Eur. J. Biochem.*, **189**:523-527,1990.

NUCLEOTIDE SEQUENCE OF THE TOXIC DOMAIN OF AN INSECTICIDAL PROTEIN GENE FROM *B. THURINGIENSIS* SUBSP. *KURSTAKI* HD-1

Qiao Liya Tian Yingchuan Mang Keqiang

(Institute of Microbiology, Academia Sinica, Beijing 100080)

Two crystal protein genes, the 5.3kb and 6.6kb class respectively, from *Bacillus thuringiensis* subsp. *kurstaki* HD-1 (*B. t* HD-1) had been cloned previously. Based on the classification system of Hofte and Whiteley, these two genes should belong to Cry I A(b) and Cry I A(c) gene type respectively. The nucleotide sequence of the toxic domain of this Cry I A(c) gene from *B. t kurstaki* HD-1 is firstly reported here and compared with that of Cry I A(c) gene from *B. t* HD-73 and Cry I A(b) gene from *B. t* HD-1.

Key words *B. thuringiensis thuringiensis* subsp. *kurstaki* HD-1 Cry I A(c) gene; Toxic domain; Nucleotide sequence