

猪瘟病毒免化弱毒株 cDNA 片段的克隆及序列分析*

李红卫¹ 刘湘涛² 李小兵¹ 韩雪清² 涂长春¹ 殷 震¹

(¹ 长春农牧大学 长春 130062)

(² 兰州兽医研究所 兰州 730046)

关键词 猪瘟病毒免化弱毒株 cDNA 克隆 序列分析

分类号 S852.65 文献标识码 A 文章编号 0001-620X(1999)06-0554-58

猪瘟是猪最重要的传染病之一,往往给养猪业造成重大经济损失。猪瘟的病原为猪瘟病毒(HCV),属黄病毒科,瘟病毒属成员,其基因组为单股正链 RNA,长度为 12.3kb,仅含有一个大的开放阅读框架,编码一个含 3898 个氨基酸残基(AA)的多聚前体蛋白^[1,2]。目前已经定位的蛋白有 5 种,即 N^{pro}、C、E0、E1 和 E2,它们均由 HCV RNA 5'端所编码,除 N^{pro}外,其它 4 种均为 HCV 的结构蛋白^[3]。N^{pro}为具有自我催化功能的蛋白水解酶,也是多聚蛋白 N 端的第一个蛋白水解酶,分子量为 23kD,C 为构成 HCV 核衣壳的核心蛋白,分子量为 14kD。E1 以 Leu-495 为 N 末端,分子量约 33kD,是 3 种糖蛋白中分子量最小的一种^[3,4]。E1 不产生中和抗体,在 HCV 感染的细胞抽提物中,E1 常与 E2 以异源二聚体的形式免疫共沉淀。E1 可能通过与 E2 形成复合抗原结构而起作用^[3,4]。5'端非编码区可能在多聚蛋白翻译调控中具有重要的作用。

猪瘟病毒免化弱毒株(HCLV)又称 C 株,是我国培育的世界公认的免疫原性和安全性最好的猪瘟疫苗株。本研究将利用 RT-PCR 技术分三段扩增 HCLV 5'端非编码区和 N^{pro}、C、E1 编码区,并对它们进行克隆和序列测定。

1 材料和方法

1.1 病毒的增殖及 RNA 的提取^[5]

HCLV 479 代冻干毒引自中国兽药监察所,在健康家兔体内增殖,采取呈定型热反应家兔的脾脏,用异硫氰酸胍一步法提取其总 RNA(采用 BRL 公司 TRI_{zol} 试剂盒),琼脂糖凝胶电泳鉴定其完整性,GeneQuant 核酸定量仪(Pharmacia 公司)测定其纯度和浓度。

1.2 引物的设计与合成

根据 HCV C 株的序列设计合成 9 条引物^[6],引物的序列及在基因组中的位置分别为:

- PA1 5' GTATACGAGGTTAGTTCATTC 3'(正链),1~21;
PA2 5' TACACCGGTTCTCCACTC 3'(负链),447~429;
PB1 5' TCTGGATCCTATAAAAACTCTGCTGTACATGG 3'(正链),347~370;
PB2 5' TTCGGCTTCAACTGGTTGGT 3'(负链),1177~1158;
PB3 5' CTTTATCGTACCTGCAAGTC 3'(负链),1529~1510;
PC1 5' GAATAGAACCCAAGCAGAC 3'(正链),1456~1474;
PC2 5' CTTATGCCCTATCGCCTTA 3'(正链),1848~1866;
PC3 5' CTGTAATCTTCCTTGCAAGC 3'(负链),2466~2447;
PC3 5' ACCACTGGCTCGCCTTTTAC 3'(负链),3024~2995。

* 国家攀登计划 B 类项目资助课题(No. 85-44-01-03)

收稿日期:1998-04-30,修回日期:1998-10-15

1.3 RT-PCR

各取上面提取的 RNA10 μ g ,分别加引物 PB2、PB3、PC4 1 μ L(200ng/ μ L) ,65 $^{\circ}$ C 5min ,立即冰浴 5min ; 然后分别依次加入 5 $\times$ RT 缓冲液 4 μ L ,dNTP(10mmol/L)1 μ L ,AMV(10u/ μ L)1 μ L ,RNasin(40u/ μ L)1 μ L , 用水补至 20 μ L ,42 $^{\circ}$ C 1h。

各取反转录产物 5 μ L ,分别加引物 PA1/PB2、PB1/PB3 和 PC1/PC4 各 1 μ L(200ng/ μ L) ,10 $\times$ PCR 缓冲液 5 μ L ,2.5mmol/L MgCl $_2$ 5 μ L(25mmol. :) dNTP(10mmol/L)1 μ L ,用水补至 49 μ L ,然后 98 $^{\circ}$ C 8min , 加 Taq DNA 聚合酶 1 μ L(1u/ μ L , Promega 公司) ,以 95 $^{\circ}$ C 50s , 48 $^{\circ}$ C 1min ,72 $^{\circ}$ C 2min ,于 Bio Techs— 97DNA Thermocycle(Beijing Biotech Co Ltd)进行循环反应。各取 PCR 产物 1 μ L 50 倍稀释 ,取 5 μ L 作为 模板 ,分别 PA1/PA2、PB1/PB2 和 PC2/PC3 为引物 ,与上面相同条件进行 PCR。

1.4 PCR 产物的克隆与序列测定

用 Agarose Gel DNA Extraction kit(Boehringer Mannheim 公司)纯化 PCR 产物 ,然后采用 pGEM—T 载体(Promega 公司)进行克隆(按操作说明书进行)。克隆后用 ABI PRISM 377 DNA Sequencer 对扩增 片段进行序列测定。

2 结果和讨论

由于 HCLV 脾毒中含病毒量很低 ,而且不易纯化 ,本研究实验开始阶段遇到了很大困难 ,单独一次 PCR 均不能扩增出特异性片段 ,后来我们采用套式或半套式 PCR 终于解决了这一难题 ,成功扩增出了 HCLV 三个与预计大小相符的 cDNA 片段 :A(PA1/PA2) ,B(PB1/PB2) ,C(PC2/PC3) ,大小应分别为 447bp、840bp、619bp(电泳图略) ,这三个片段分别包括 5'端非编码区、N^{PR}(p23)和 C(p14)编码区、E1 基 因。用套式或半套式 PCR 扩增 HCLV 基因组的其它区域也获得了成功 ,可见该方法对扩增低拷贝基因 具有明显的优势。

将扩增片段克隆到 pGEM-T 载体中 ,用 PCR 鉴定重组质粒(电泳结果略) ,然后用全自动序列分析 仪对插入片段进行序列测定 ,结果见图 1、2、3。将所测定的部分序列(包括部分 5'端非编码区、p23 和部 分 214 基因)与 HCV 石门株(我国 HCV 标准强毒株)及 C 株(Moormann 等测得)相应序列进行比较(见 图 4) ,结果表明 HCLV 与石门株及 C 株的同源性分别为 94.69% 和 99.19%。在序列比较中我们还发 现 ,C 株与 HCLV 和石门株相比缺乏 244 处的碱基 C ,但因为该碱基位于 5'端非编码区 ,因此并不引起 读码框架的改变。本实验的序列比较结果表明 HCLV 和 Moormann 等用于测序的 C 株来源可能不同。

```
1 TCGTGTACAC GATTGGACAA ATCACAATTA TAATTTGGTT CAGGGCCTCC CTCCAGCGAC
61 GGCCGAAC TG GGCTAGCCAT GCCCATAGTA GGACTAGCAA AACGGAGGGA CTAGCCATAG
121 TGGCGAGCTC CCTGGGTGGT CTAAGTCTCT AGTACAGGAC AGTCGTCAGT AGTTCGACGT
181 GAGCAGAAGC CCACCTCGAG ATGCTACGTG GACGAGGGCA TGCCCAAGAC ACACCTTAAC
241 CCTAGCGGGG GTCGCTAGGG TGAAATCACA CCACGTGATG GGAGTGGCAG CTGATAGGGC
301 GCTGCAGAGG CCCACTATTA GGCTAGTATA AAAATCTCTG CTGTACATGG CACATGGAGT
361 TGAATCACTT TGAAC TTTTA TACAAAACAA ACAACAAAA ACCAATGG
```

图 1 测定的 HCLV A 片段(可能包括完整的 5'端非编码区)序列

本研究首次完成了我国 HCLV 的 5'端非编码区及 p23、p14、E1 编码区的克隆及序列测定 ,这将对 进一步研究这些基因区域或编码蛋白的功能打下了基础 ,也可能为猪瘟的防制提供新的途径。

1	CAC	ATG	GAG	TTG	AAT	CAC	TTT	GAA	CTT	TTA	TAC	AAA	ACA	AAC	AAA	CAA	AAA	CCA	ATG
1		M	E	L	N	H	F	E	L	L	Y	K	T	N	K	Q	K	P	M
58	GGA	GTG	GAG	GAA	CCG	GTG	TAC	GAT	GCC	ACG	GGG	AGA	CCA	TTG	TTC	GGA	GAC	CCG	AGT
19	G	V	E	E	P	V	Y	D	A	T	G	R	P	L	F	G	D	P	S
115	GAG	GTA	CAC	CCA	CAA	TCA	ACA	CTA	AAG	CTA	CCA	CAT	GAT	AGG	GGT	AGA	GGC	AAC	ATT
38	E	V	H	P	Q	S	T	L	K	L	P	H	D	R	G	R	G	N	I
172	AAA	ACA	ACA	CTG	GAG	AAC	CTA	CCT	AGG	AAA	GGC	GAC	TGC	AGG	AGC	GGC	AAC	CAT	CTA
57	K	T	T	L	E	N	L	P	R	K	G	D	C	R	S	G	N	H	L
229	GGC	CCG	GTC	AGT	GGG	ATA	TAT	GTA	AAA	CCC	GGC	CCT	GTC	TTT	TAC	CAG	GAC	TAC	ATG
76	G	P	V	S	G	I	Y	V	K	P	G	P	V	F	Y	Q	D	Y	M
286	GGC	CCG	GTC	TAC	CAT	AGA	GCC	CCT	CTG	GAG	TTT	TTT	GAC	GAA	GTG	CAG	TTC	TGC	GAG
95	G	P	V	Y	H	R	A	P	L	E	F	F	D	E	V	Q	F	C	E
343	GTG	ACC	AAA	AGG	ATA	GGT	AGG	GTG	ACG	GGT	AGC	GGA	GGA	AAG	CTT	TAC	CAT	ACA	TAT
114	V	T	K	R	I	G	R	V	T	G	S	D	G	K	L	Y	H	T	Y
400	GTG	TGC	ATC	GAT	GGC	TGC	ATA	CTG	CTG	AAG	CTG	GCC	AAG	AGG	GGT	GAG	CCA	AGA	ACC
133	V	C	I	D	G	C	I	L	L	K	L	A	K	R	G	E	P	R	T
457	CTG	AGG	TGG	ATT	AGA	AAT	TTC	ACC	GAC	TGT	CCA	TTG	TGG	GTT	ACC	AGT	TGC	TCC	GAT
152	L	R	W	I	R	N	F	T	D	C	P	L	W	V	T	S	C	S	D
514	GAT	GGC	GCA	AGT	GGG	AGT	AAA	GAG	AAG	AAG	CCA	GAT	AGG	ATC	AAT	AAA	GGC	AAA	TTA
171	D	G	A	S	G	S	K	E	K	K	P	D	R	I	N	K	G	K	L
571	AAA	ATA	GCC	CCA	AAA	GAG	CAT	GAG	AAG	GAC	AGC	AGA	ACT	AGG	CCA	CCT	GAC	GCT	ACG
190	K	I	A	P	K	E	H	E	K	D	S	R	T	R	P	P	D	A	T
628	ATC	GTG	GTA	GAA	GGA	GTA	AAA	TAC	CAG	GTC	AAA	AAG	AAA	GGT	AAA	GTT	AAA	GGG	AAG
209	I	V	V	E	G	V	K	Y	Q	V	K	K	K	G	K	V	K	G	K
685	AAT	ACC	CAA	GAC	GGC	CTG	TAC	CAC	AAC	AAG	AAT	AAA	CCA	CCA	GAA	TCC	AGG	AAG	AAA
228	N	T	Q	D	G	L	Y	H	N	K	N	K	P	P	E	S	R	K	K
742	TTA	GAA	AAA	GCC	CTA	TTG	GCA	TGG	GCG	GTG	ATA	GCA	ATT	ATG	TTG	T			
247	L	E	K	A	L	L	A	W	A	V	I	A	I	M	L				

图2 测定的HCLV B片段的序列(包括完整的 p23 和 p14 基因)和诱导的氨基酸序列

1	C	TGT	AAT	GTA	ACA	AGC	AAG	ATA	GGG	TAC	ATA	TGG	TAC	ACT	AAC	AAC	TGC	ACC	CCA
1		C	N	V	T	S	K	I	G	Y	I	W	Y	T	N	N	C	T	P
56	GCT	TGC	CTC	CCC	AAA	AAC	ACA	AAG	ATA	ATA	GGC	CCT	GGT	AAA	TTT	GAC	ACT	AAT	GCA
19	A	C	L	P	K	N	T	K	I	I	G	P	G	K	F	D	T	N	A
113	GAA	GAC	GGA	AAG	ATT	CTC	CAT	GAG	ATG	GGG	GGC	CAC	CTA	TCA	GAA	TTT	CTG	CTG	CTT
38	E	D	G	K	I	L	H	E	M	G	G	H	L	S	E	F	T	L	L
170	TCT	CTG	GTT	GTT	CTG	TCT	GAC	TTC	GCC	CCA	GAA	ACA	GCC	AGC	GCG	TTA	TAC	CTC	ATT
57	S	L	V	V	L	S	D	F	A	P	E	T	A	S	A	L	Y	L	I
227	TCG	CAC	TAC	GTG	ATT	CCT	CAA	CCC	CAT	GAT	GAA	CCT	GAA	GGC	TGC	GAT	ACG	AAC	CAG
76	S	H	Y	V	I	P	Q	P	H	D	E	P	E	G	C	D	T	N	Q
284	CTG	AAT	CTA	ACA	GTA	GAA	CTC	AGG	ACT	GAA	GAC	GTA	ATA	CCG	TCA	GTC	TGC	TGG	AAT
95	L	N	L	T	V	E	L	R	T	E	D	V	I	P	S	S	V	W	N
341	GTT	GGT	GAA	TAT	GTG	TGT	ATT	AGA	CCA	GAC	TGG	TGG	CCA	TAT	GAA	ACC	GAG	GTG	GCT
114	V	G	E	Y	V	C	I	R	P	D	W	W	P	Y	E	T	E	V	A
398	CTG	TTA	TTT	GAA	GAG	GCA	GGA	CAG	GTC	GTA	AAG	TTA	GTC	TTA	CGG	GCG	CTG	AGG	AAT
133	L	L	F	E	E	A	G	Q	V	V	K	L	V	L	R	A	L	R	D
455	TTG	ACT	AGG	GTC	TGG	AAT	AGC	GCA	TCA	ACC	ATT	GCA	TTC	CTC	ATC	TGC	TTG	ATA	AAA
152	L	T	R	V	W	N	S	A	S	T	I	A	F	L	I	C	L	I	K
512	GTA	TTA	AGG	GGA	CAG	ATC	GTG	CAA	GGT	GTG	GTA	TGG	CTG	TTA	CTA	GTA	ACT	GGG	GCA
171	V	L	R	G	Q	I	V	Q	G	V	V	W	L	L	L	V	T	G	A
569	CAA	GGC	CGG	CTA															
190	Q	G	R	L															
.....	..	..	..	..															

图3 测定的HCLV C片段序列(包括完整的 E1 基因)和诱导的氨基酸序列

HCLV. DNA	1	CGAGCTCCCT	GGGTGGTCTA	AGTCCTGAGT	ACAGGACAGT	CGTCAGTAGT
SHIMEN. DNA	1	*****	*****	*****	*****	*****
C. DNA	1	*****	*****	*****	*****	*****
HCLV. DNA	51	TCGACGTGAC	CAGAAGCCCA	CCTCGAGATG	CTACGTGGAC	GAGGGCATGC
SHIMEN. DNA	51	*****	*****	*****	*****	*****
C. DNA	51	*****	*****	*****	*****	*****
HCLV. DNA	101	CCAAGACACA	CCTTAACCCT	AGCGGGGGTC	GCTAGGGTGA	AATCACACCA
SHIMEN. DNA	101	*****	*****	*****	*****	*****
C. DNA	101	*****	*****	*****	*****	*****
HCLV. DNA	151	CGTGATGGGA	GTGCGACCTG	ATAGGGCGCT	GCAGAGGCC	ACTATTAGGC
SHIMEN. DNA	151	*****	*****	*****	*****	*****
C. DNA	151	*****	*****	*****	*****	*****
HCLV. DNA	201	TAGTATAAAA	ATCTCTGCTG	TACATGGCAC	ATGGAGTTGA	ATCACTTTGA
SHIMEN. DNA	201	*****	*****	*****	*****	*****
C. DNA	201	*****	*****	*****	*****	*****
HCLV. DNA	251	ACTTTTATAC	AAAACAAACA	AACAAAACC	AATGGGAGTG	GAGGAACCGG
SHIMEN. DNA	251	*****	*****	*****	*****	*****
C. DNA	251	*****	*****	*****	*****	*****
HCLV. DNA	301	TGTACGATGC	CACGGGGAGA	CCATTGTTGC	GAGACCCGAG	TGAGGTACAC
SHIMEN. DNA	301	*****	*****	*****	*****	*****
C. DNA	301	*****	*****	*****	*****	*****
HCLV. DNA	351	CCACAATCAA	CACCTGAAGCT	ACCACATGAT	AGGGGTAGAG	GCAACATTAA
SHIMEN. DNA	351	*****	*****	*****	*****	*****
C. DNA	351	*****	*****	*****	*****	*****
HCLV. DNA	401	AACAACACTG	GAGAACCTAC	CTAGGAAAGG	CGACTGCAGG	AGCGGCAACC
SHIMEN. DNA	401	*****	*****	*****	*****	*****
C. DNA	401	*****	*****	*****	*****	*****
HCLV. DNA	451	ATCTAGGCC	GGTCAGTGGG	ATATATGTAA	AACCCGGCCC	TGCTTTTAC
SHIMEN. DNA	451	*****	*****	*****	*****	*****
C. DNA	451	*****	*****	*****	*****	*****
HCLV. DNA	501	CAGGACTACA	TGGGCCCCGT	CTACCATAGA	GCCCCCTCTG	AGTTTTTTGA
SHIMEN. DNA	501	*****	*****	*****	*****	*****
C. DNA	501	*****	*****	*****	*****	*****
HCLV. DNA	551	CGAAGTGCAG	TTCTGCGAGG	TGACCAAAAG	GATAGGTAGG	GTGACGGGTA
SHIMEN. DNA	551	*****	*****	*****	*****	*****
C. DNA	551	*****	*****	*****	*****	*****
HCLV. DNA	601	GCGACGGAAA	GCTTTAACCAT	ACATATGTGT	GCATCGATGG	CTGCATACTG
SHIMEN. DNA	601	*****	*****	*****	*****	*****
C. DNA	601	*****	*****	*****	*****	*****
HCLV. DNA	651	CTGAAGCTGG	CCAAGAGGGG	TGAGCCAAGA	ACCCTGAGGT	GGATTAGAAA
SHIMEN. DNA	651	*****	*****	*****	*****	*****
C. DNA	651	*****	*****	*****	*****	*****
HCLV. DNA	701	TTTCACCGAC	TGTCCATTGT	GGGTTACCAG	TTGCTCCGAT	GATGGCGCAA
SHIMEN. DNA	701	*****	*****	*****	*****	*****
C. DNA	701	*****	*****	*****	*****	*****
HCLV. DNA	751	GTGGGAGTAA	AGAGAAGAAG	CCAGATAGGA	TCAATAAAGG	CAAATTAATA
SHIMEN. DNA	751	*****	*****	*****	*****	*****
C. DNA	751	*****	*****	*****	*****	*****
HCLV. DNA	801	ATAGCCCCAA	AAGAGCATGA	GAAGGACAGC	AGAACTAGGC	CACCTGACGC
SHIMEN. DNA	801	*****	*****	*****	*****	*****
C. DNA	801	*****	*****	*****	*****	*****
HCLV. DNA	851	TACGATCGTG	GTGGAA. . .			
SHIMEN. DNA	851	*****	*****			
C. DNA	851	*****	*****			

图 4 HCV 不同来源 C 株及石门株 (SHIMEN. DNA) 之间部分基因组序列的比较
HCLV. DNA 来自本研究, C. DNA 来自 Moormann 等测得的 C 株序列, 比较的序列包括部分 5' 端非编码区、p23 基因和部分 p14 基因, 第 1 个碱基对应于 C 株基因组的 145 位碱基。

参 考 文 献

- [1] Meyers G, Rumenapf T, Thiel H J. *Virology*, 1989, **171**(2):555~567.
- [2] Moormann R J M, Warmerdam P A M, Meer B V D *et al.* *Virology*, 1990, **177**(1):184~198.
- [3] Stark R, Rumenapf T, Meyers G *et al.* *Virology*, 1998, **174**(1):286~289.
- [4] Thiel H J, Stark R, Weiland E *et al.* *J Virol*, 1991, **65**(9):4705~4712.
- [5] Chomezynski P, Piot, Nicoletta S. *Analytical Biochemistry*, 1987, **162**(1):156~159.
- [6] Moormann, R J, van Gennip H g, Miedema G K *et al.* *J Virol*, 1996, **70**(2):763~770.

MOLECULAR CLONE AND SEQUENCE ANALYSIS OF cDNA FRAGMENTS OF HOG CHOLERA VIRUS STRAIN C*

Li Hongwei¹ Liu Xiangtao² Li Xiaobing¹ Han Xueqing² Tu Changchun¹ Yin Zhen¹

(¹ Changchun University of Agricultural and Animal Sciences, Chnagchun 130062)

(² Lanzhou Veterinary Institute, Lanzhou 730046)

Abstract The 5' nonencoding region, p23 and p14 encoding region and E1 gene of hog cholera virus (HCV) strain C were amplified from total RNA extracted from HCV strain C infected rabbit spleen by reverse transcription and nested or half-nested PCR. The PCR products were cloned into pGEM-T vector. Nucleotide sequencing was performed using an ABI PRISM sequencing device; based on the incorporation of fluorescein labelled dideoxynucleotide terminators. The obtained sequences on 5' noncoding region and part of p23 and p14 encoding region were compared with HCV strains Shimen and C sequenced in Moormann's lab. The result showed that the homology between HCV strains C sequenced in this report and in Moormann's lab was 99.19%, and the homology between HCV strains Shimen, the standard virulent HCV strain in China, and C sequenced in this report was 94.69%. It was also discovered that the base C at 244 of the genome of HCV strains Shimen and C sequenced in this report was absent at the genome of C strain sequenced in Moormann's lab *et al.*

Key words Hog cholera virus strain C, cDNA fragments, Cloning, Sequencing