

## HCV E2 区基因的分子克隆及序列分析\*

张东伟<sup>1</sup>, 梁布锋<sup>1\*\*</sup>, 祁自柏<sup>2</sup>, 凌世淦<sup>3</sup>(1 中国科学院武汉病毒研究所, 湖北武汉 430071; 2. 中国药品生物制品检定所, 北京 100050;  
3. 军事医学科学院, 北京 100850)

## Molecular Cloning and Sequencing of HCV E2 Gene

ZHANG Dong-wei<sup>1</sup>, LIANG Bu-feng<sup>1\*\*</sup>, QI Zi-bai<sup>2</sup>, LING Shi-gan<sup>3</sup>(1 *Wuhan Institute of Virology, Academia Sinica, Wuhan 430071, China;*  
2. *Dept. of Hepatitis, National Institute for the Control of Pharmaceutical  
and Biology Products, Beijing 100050, China;*  
3. *Military Medical Academia Sinica, Beijing 100850, China)*

**Abstract:** HCV RNA positive serum was first selected by RT-PCR test kit from several anti-HCV positive sera obtained from Xi'an. HCV RNA extracted from the elected sera was converted to cDNA by reverse transcription with random primer. Half-nested PCR was performed. The amplified product was 852 bp. The purified PCR product was digested by restriction endonucleases and then ligated to episio vector pET-22b<sup>+</sup>. Its nucleotide sequence was determined by dideoxy chain termination method. A comparison of the sequence with several isolates reported previously showed that the sequence belonged to HCV type II.

**Key words:** Hepatitis C Virus; E2 glycoprotein; HVR1; Sequence character

**摘要:** 用 RT-PCR KIT 从西安血站大样抗 HCV 阳性血清中筛选出 HCV RNA 阳性血清, 提取 HCV 的 RNA, 利用随机引物反转录合成其 cDNA 并进行半巢式 PCR 反应。将纯化的 PCR 产物酶切后与表达载体 PET-22b<sup>+</sup> 连接, 经过双脱氧末端终止法双向测序, 得到 852 bp 长的核苷酸序列。通过将该序列与已知不同型的 HCV E2 序列比较得知, 此序列正是 HCV II 型目的基因。

**关键词:** 丙型肝炎病毒; E2 包膜糖蛋白; HVR1; 序列特征

**中图分类号:** Q785 **文献标识码:** A **文章编号:** 1003-5125(2001)01-0040-05

丙型肝炎病毒(HCV)是肠道外非甲非乙型肝炎的主要病原, 目前全世界约有 1.7 亿 HCV 感染者, 我国预计受感染者在 2 千万以上。由于目前还没有确实有效的治疗方法, 感染者中的大部分都导致成慢性肝炎, 进而部分可发展成肝硬化和肝细胞癌。HCV 属黄病毒科, 为单股正链 RNA 病毒, 全长约有 9.5 kb, 编码 3 010~3 033 个氨基酸, 蛋白排列

依次为 5'NCR-C-E1-E2-NS2-NS5-3'NCR。其中 E2 区是 HCV 的包膜糖蛋白基因, 编码 384 aa~746 aa 的一段蛋白<sup>[1]</sup>。E2 基因上有两个高变区 HVR1 (384 aa~410 aa)和 HVR2(475 aa~482 aa)。研究已证明 HVR1 上有 B 细胞识别的抗原表位, 可诱导产生中和性抗体。但由于 HVR1 的多变性, 使得病毒逃避宿主的免疫压力而在机体内持续存在, 这可

收稿日期: 2000-04-21, 修回日期: 2000-07-18

\* 基金项目: 国家九五科技攻关基金资助项目(96-90-603-06)

作者简介: 张东伟(1969-), 男, 陕西省籍, 硕士研究生, 主要从事医学病毒的分子生物学研究。

\*\* 通讯作者: 梁布锋(1947-), 男, 广东省籍, 研究员。For correspondence.

能正是 HCV 感染易慢性化的原因<sup>[2]</sup>。

本文采用 RT-PCR 的方法从病人血清中扩增到 HCV II 型的编码从 376 aa~659 aa 的一段基因,将其连接到表达载体上,这为以后 E2 蛋白的表达奠定基础。同时将该序列与不同型的 HVR1 的核苷酸序列及氨基酸序列分别进行了比较,以找出其变化规律。

## 1 材料与方法

### 1.1 血清样本

5 份抗 HCV 阳性血清,取自西安血站大样。

### 1.2 主要试剂

RT-PCR 检测试剂盒购自 302 医院;RNA 提取试剂购自 BIONEER 公司;AMV、RNasin、Taq 酶、*Bam*HI、*Hind*III 限制性内切酶、随机引物、琼脂糖及 T4 连接酶均购自 Promega 公司;DEPC 购自 Sigma 公司;PCR 产物纯化试剂盒为上海华舜生物公司产品;PCR 仪为 PE 公司产品;载体 pET-22b<sup>+</sup>、大肠杆菌 DH5 $\alpha$  为本室保存。

### 1.3 引物设计

主要参照 HCV-HB、HPC1b1、HPC1b2 等 HCV II 型 E2 的保守区设计,由上海生工生物公司合成。序列如下:

外引物:上 1a 5'1126-CCUAAGCTTCTACTCTTSGCCGGGTTG

下 2b 5'2238-AATTGGGATCCGGTCCCTCYGGCTGYGCTA

内引物:上 2a 5'1126-ATAGCATUCACTACTCTTSGCCGGGTTG

下 2b 5'1977-CCCAAGCTTTTCATCCTGTCTCCA

鉴定内引物:3b 5'1467-TACGGGATCCGGTAGTGCCAGCAATAGG

注:S=T/C Y=A/G

### 1.4 HCV RNA 阳性血清的筛选

用 HCV RT-RNA 检测试剂盒从抗 HCV 阳性的血清中筛选出 HCV RNA 阳性血清。

### 1.5 cDNA 产物合成

取筛选出的阳性血清 200  $\mu$ L,按照 RNA 提取说明书操作,获得的 RNA 溶在 40  $\mu$ L 水中,取 2  $\mu$ L 为模板,加入 0.6  $\mu$ L 随机引物(0.5  $\mu$ g/ $\mu$ L),再加入 AMV 5  $\times$  buffer 5  $\mu$ L、dNTP 2.5  $\mu$ L、RNasin 40 units、AMV-RT 12 units,加 DEPC 水至 15  $\mu$ L,70 $^{\circ}$ C 变性 5 min,立刻在冰上放置 5 min,离心几秒钟,再加入 DEPC 水至 25  $\mu$ L,37 $^{\circ}$ C 孵育 1 h 进行 cDNA 合成。

### 1.6 PCR 产物的获得

第一次 PCR:取逆转录产物 5  $\mu$ L 为模板,加入 10  $\times$  buffer、dNTP、MgCl<sub>2</sub>、外引物 1a 和 1b、Taq 酶,总体积 25  $\mu$ L,94 $^{\circ}$ C 预变性 5 min,然后 94 $^{\circ}$ C 60 s,

50 $^{\circ}$ C 90 s,72 $^{\circ}$ C 150 s 共 30 个循环,再 72 $^{\circ}$ C 延伸 7 min;第二次 PCR:取第一次 PCR 产物 2.5  $\mu$ L 为模板,加入内引物 2a、2b,反应条件同上,总体积 25  $\mu$ L,用双温 PCR 扩增:

94 $^{\circ}$ C 60 s,50 $^{\circ}$ C 90 s,72 $^{\circ}$ C 150 s 2 个循环

94 $^{\circ}$ C 60 s,55 $^{\circ}$ C 90 s,72 $^{\circ}$ C 150 s 28 个循环

再 72 $^{\circ}$ C 延伸 7 min,取 10  $\mu$ L 在 1.2% 的琼脂糖凝胶中进行电泳,紫外灯下观察结果。

### 1.7 E2 基因的克隆及鉴定

PCR 产物用纯化柱纯化后,利用 *Bam*HI、*Hind*III 进行双酶切,再用纯化柱纯化,与同样处理的载体 pET-22b<sup>+</sup> 在 14~16 $^{\circ}$ C 连接 14 h,连接反应总体积为 20  $\mu$ L;取约 10  $\mu$ L 转化 DH5 $\alpha$  感受态细胞。碱裂解法提质粒,用 *Bam*HI、*Hind*III 双酶切鉴定,再用 *Not*I 单酶切鉴定,最后用引物 2a 和 3b 进行半巢式 PCR 鉴定。

### 1.8 DNA 序列测定

以重组的 pET 质粒(pET-E2)为模板,采用双脱氧末端中止法在 ABI PRISM 377XL DNA 全自动测序仪上双向测序。

## 2 结果

### 2.1 阳性克隆 pET-E2 的鉴定

E2 的 PCR 酶切产物连接到 pET-22b<sup>+</sup> 的多克隆位点的 *Bam*HI 和 *Hin*dIII 之间,用这两种酶切

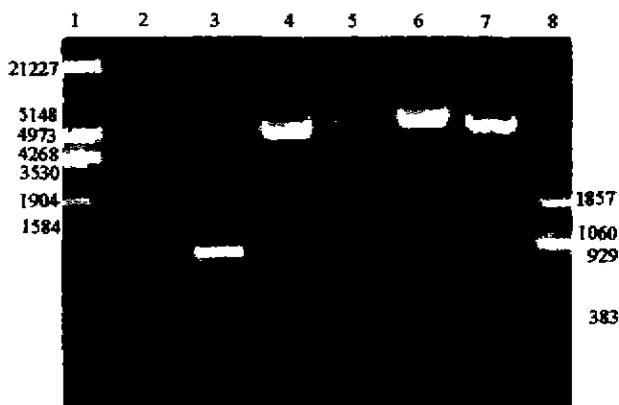


图 1 PCR 和酶切鉴定重组质粒 pET-E2  
Fig.1. Recombinant plasmid pET-E2 identified by PCR and restriction endonucleases

1, DNA molecular weight marker ( $\lambda$  DNA/*Eco*R I & *Hin*d III); 2, PCR products of pET-E2 amplified with primers 2a and 3b; 3, RT-PCR products amplified with primers 2a and 2b from positive serum; 4, pET-E2 digested by *Bam*HI and *Hin*d III; 5, pET-22b digested by *Bam*HI; 6, pET-E2 digested by *Bam*HI; 7, pET-E2 digested by *Not*I; 8, DNA molecular weight marker (pBR322/*Bst*NI).

pET-E2 就得到两条带, 分别与 pET-22b<sup>+</sup> 单切带和 E2 PCR 产物带相当。由于引物是按照 HCV II 型保守区设计, HCV II 型在 E2 的第 500~507 位碱基正是 *NotI* 的切点, 同时 pET-E2 下游的多克隆位点上也有 *NotI* 的切点, 因此, *NotI* 单切即得到 350 bp 的

一条带; 鉴定内引物 3b 也按照内引物 2a、2b 之间的保守区设计, 以 2a、3b 为引物做 PCR 得到 363 bp 的一条带, 此作用相当于 DNA 探针杂交, 进一步证实了所得到的目的基因 HCV II 型 E2 (见图 1), 此株病毒且称之为 HCV-XA。

**CTA CTC TTC GCC GGC GTT GAT** GGG GCG ACC CAT GTG ACG GGG GGG 45  
 L L F A G V D G A T H V T G G  
 GTG GCA GGC AGA ACC ACC CAT GGG TAC GCG TCC TTG TTT TCG GTT 90  
 V A G R T T H G Y A S L F S V  
 GGG GCG TCT CAG AAG CTC CAG CTT GTA AAC ACC AAT GGC AGT TGG 135  
 G A S Q K L Q L V N T N G S W  
 CAC ATC AAC AGG ACA GCC CTG AAC TGC AAT GAC TCC CTC AAC ACT 180  
 H I N R T A L N C N D S L N T  
 GGG TTC CTT GCT GCG CTG TTC TAC CAC CAC AAG TTC AAC TCG TCC 225  
 G F L A A L F Y H H K F N S S  
 GGG TGT ATA GAG CGC ATG GCC AGC TGC CGC TCC ATT GAC GAG TTT 270  
 G C I E R M A S C R S I D E F  
 GCT CAG GGA TGG GGT CCT ATC ACT TAT GCT CAG CCT GGC AAC CTG 315  
 A Q G W G P I T Y A Q P G N L  
 GAC CAG AGG CCT TAT TGC TGG CAC TAC TCA CCT CGG CCG TGT GGC 360  
 D Q R P Y C W H Y S P R P C G  
 ATC GTG CCC GCG TCA GAG GTG TGT GGT CCA GTG TAT TGT TTC ACC 405  
 I V P A S E V C G P V Y C F T  
 CCT AGC CCT GTT GTG GTG GGG ACG ACC GAC CGC TCT GGT GTT CCC 450  
 P S P V V V G T T D R S G V P  
 ACG TAT AGC TGG GGG GAG AAT GAG ACA GAC GTG CTG CTC CTC AAC 495  
 T Y S W G E N E T D V L L L N  
 AAC ACG CGG CCG CCG CAA GGC AAC TGG TTC GGC TGT ACA TGG ATG 540  
 N T R P P Q G N W F G C T W M  
 AAT GCT ACC GGG TTC ACC AAG ACG TGC GGG GGC CCC CCG TGC AAC 585  
 N A T G F T K T C G G P P C N  
 ATC GGG GGG GTC GGC AAC AAC ACC TTG ATT TGC CCC ACG GAC TGC 630  
 I G G V G N N T L I C P T D C  
 TTC CGG AAG CAC CCC GAG GCC ACT TAC ACA AAA TGT GGT TCT GGG 675  
 F R K H P E A T Y T K C G S G  
 CCT TGG TTA ACA CCT AGG TGC TTA GTC GAC TAT CCA TAC AGG CTI 720  
 P W L T P R C L V D Y P Y R L  
 TGG CAC TAC CCC TGC ACT GTC AAT TTC ACC GTC TTC AAG GTT AGG 765  
 W H Y P C T V N F T V F K V R  
 ATG TAT GTG GGG GGC GTG GAG CAC AGG CTC AAT GCT GCG TGC AAT 810  
 M Y V G G V E H R L N A A C N  
 TGG ACT CGA GGA GAG CGT TGT AAC **TG GAG GAC AGG GAT AGA** 855  
 W T R G E R C N L E D R D R

图2 E2的部分片段的cDNA序列和氨基酸序列

Fig.2 cDNA sequence and amino acid sequence of E2

2.2 HCV 西安株 E2 区的 cDNA 序列和氨基酸序列

HCV-XA cDNA 序列和氨基酸序列见图 2(斜体为引物部分), 通过 GOLDKEY 软件分析, HCV-XA 与其它 HCV II 型序列的同源性在 85% 以上, 而

与其它型的同源性均小于 73% (比较结果未列出)。

2.3 HCV 西安株的 HVR1 与其他已知序列的比较

应用计算机分析软件将所测得的序列与其它已知序列进行同源性比较, 其结果见图 3、图 4。

STRAINS	SEQUENCES	IDENTITY	SCORE
E2-661	GGCACCCATGTGACGGGGGGTGGCAGGCAGAACCCACCCATGGGTA CGCGTCTTGTTTTCGGTTGGGCGTCTCAGAAG		
Hpc1b2	AGC. C T.A. . . . . CACTA. . CTTAC. AT. GGA. C T.TA. . . . . C.C. . . . . ATC. . . . . CAA. . . . . A	58.02%	47
Hpc1b3	. AC T.CAC. . A. . . . . C CAG CAC . TAG. A. GT.G A C.C. . A ATC. . . A. A. . . . . T	61.73%	50
Hpc1b4	AGT GC T.A. . . . . CACTA. CTTAC. AT GG. . CCT.TA C.C. . T ATC. . . . . CAA. . . . . A	56.79%	46
Hpc1b5	CGC ACC. . A. . . . . C CAG CAT . . . . . AA. . CT . . . . . C.C. . . A. AT. . . . . CAG. . . . . GA	65.43%	53
Hpcy1b6	AAC T.CAC A. . . . . C C CAG C.C. . GTTAG A. T.GA. A A.C. . . A ACC. . . . . A. A. . . . . T	58.02%	47
HCV-HB	. AT T.CGC.T.T. . . . . C CAG C.CT. . . . . TC. . . T.A. . . C.C. . A ACC. . . . . C. . . . .	69.14%	56
HCV-T	. TAC ATCGTGT. A. . A. ACA. TG. C. C. C. . . . . CA. TCT . . . . . TC. C. . A. ACAA. . . . . C. . . . . A	55.56%	45
HPCGGG	CAC. . C. . . . . A. . . . . AAG. . T. . C TCC. G. . . . . GA. CCT . T. . . . . G. C. C. ACAA. . CC. A. . . . . A	61.73%	50
HCV-1	. AA. . . C. . C. . C. . . . . AAGT. . C. . CAC . TGTGTC. . . A. T. T. TTAG C. CC. CG. ACCA. . C. . CAAG. . . . . C	49.38%	40

图 3 HCV 西安株的 HVR1 cDNA 与其它株相应序列的比较

其它株序列见文献: Hpc162<sup>[3]</sup>, Hpc163<sup>[3]</sup>, Hpc164<sup>[3]</sup>, Hpc165<sup>[3]</sup>, Hpc166<sup>[3]</sup>, HCV-HB<sup>[4]</sup>, HCV-T<sup>[3]</sup>, HCV-1<sup>[6]</sup>。

Fig. 3 Comparison of cDNA sequence of HCV-XA HVR1 with corresponding sequences of other HCV strains.

STRAINS:	SEQUENCES	IDENTITY
HCV-XA	ATHVTGGVAGRITTHGYASLFSVGSQK	
HPC1B2	S---S--JTAYN-NSFT---S-Q--	51.85%
HPC1B3	D-YT---AQ-H-SRV---TS---N	55.56%
HPC1B4	S-R-S--JTAYN-R-LT---S-Q--	51.85%
HPC1B5	R-T---AQ-H-QRL---TF-Q-R	55.56%
HPCY1B6	N-YT--AAQ--VSRLT-I-TP---N	44.44%
HCV-HB	D-YAS-AQ-S-L-FT---TP---	55.56%
HCV-T	VPSCS-TVA---SL---TQ---	55.56%
HPCGGG	H---RVASS-QSLV-RL-Q-P---	50.00%
HCV-1	E---S-H-VS-FV-LAP-K-N	55.56%

图 4 HCV 西安株的 HVR1 推导的氨基酸序列与其他已知相应序列的比较

Fig. 4 Comparison of amino acid sequence of HCV-XA HVR1 with corresponding sequences of other reported strains

3 讨论

由于 HCV E2 的高变性, 我们应用 GOLDKEY 软件设计引物时, 主要参考了 Genebank 中及我们自

己输入的 HCV II 型 E2 序列的相对保守区, 在引物中使用了兼并密码子。由于血清中 HCV 浓度极低, 在扩增时我们采用双温 PCR, 这样可以在前两个循环增大模板量, 提高检出率; 扩增产物设计为 852 bp, 编码 284 个氨基酸 (376 aa~659 aa), 而不是全长, 因为 E2 蛋白 C 端的一段氨基酸序列是疏水区, 去除后可获得分泌型表达, 尤其在哺乳动物细胞中的表达产物将呈现天然构像, 有较高的生物活性<sup>[7]</sup>。

HCV 的高变性为疫苗的研制带来极大的困难。用较保守的 C 区作疫苗则不安全, 因为有研究证明, C 蛋白可能正是 HCV 的致癌因素之一<sup>[8]</sup>, 不适合作疫苗。E2 包膜糖蛋白上有两个 HCV 中和抗体的抗原表位, 一个高变, 一个保守, 进一步的研究证明, 其中一个在 HVR1 上<sup>[9]</sup>。从我们克隆的 HCV-XA 的 HVR1 与其他 II 型 HVR1 的氨基酸及核苷酸序列比较来看, HVR1 的突变多为有义突变, HCV-XA 与 HPCY1b6 的核苷酸同源性 58%, 而氨基酸同源性仅 44.44%。从氨基酸水平来看, 不同株甚至不同型的 HVR1 的同源性相差较小, HCV-XA 与 HCV-1 核苷同源性仅 49.38%, 氨基酸同源性却有 55.56%, 预示着 HVR1 变异的随机性, 即同型之间的同源性不同型之间的同源性差别不大。从图 4 来看, 在 HVR1 的 27 个氨基酸中, 有几个位

置是高度保守的:第6位G、第18位S、第23位G、第26位Q,有几个位置是比较保守的:第2位T、第7位G、第13位T、第18位S、第20位F、第24位A,只从HCVII型考虑,可以得出以下的HVR1保守区框架:

\* T \* \* (T/S)G(G/A) \* \* \* \* \* (T/V) \* \*  
\* \* SLF(S/T) \* GA \* Q \*

这表明HCVII型的HVR1上有12个保守的氨基酸残基,占总数的44%,其中,T、S、G、A、V都是小分子氨基酸,S/T都带一羟基,T/V有相似构像,G/A都可以绕轴旋转,在某种程度上可以互换而不影响蛋白的性质。这些保守位点可能是病毒与宿主细胞结合所必需的,这对以后疫苗的研制可能有积极意义。

#### 参考文献

- [1] Mizushima H, Hijikata M, Aasbe SI, *et al.* Two hepatitis C virus glycoprotein E2 products with difference C termini [J]. *J Virol*, 1994, 68:6215-6222.
- [2] Weiner A J, Geysen H M, Christopherson, *et al.* Evidence for immune election of hepatitis C virus(HCV) putative envelope glycoprotein variants: Potential role in chronic HCV infectious [J]. *Proc Natl Acad Sci*, 1992, 89:3468-3472.
- [3] Genebank 中标准 HCV II 型序列[DB].
- [4] 毕胜利、白亮鹤、丛勉尔,等. 中国人丙型肝炎病毒基因组的一级结构及变异[J]. *病毒学报*, 1993, 9:114-127.
- [5] Chen P J, Lin M H, Tal K F, *et al.* The Taiwanese hepatitis C virus genome; Sequence determination and mapping the 5' termini of viral genomic and antigenomic RNA[J]. *Virology*, 1992, 188:102-113.
- [6] Choo Q L, Richman, K H, Han J K, *et al.* Genetic organization and diversity of the hepatitis C virus[J]. *Proc Natl Acad Sci*, 1991, 88:2451-2455.
- [7] Michalak J P, Wychowski C, Choukhi A, *et al.* Characterization of truncated forms of hepatitis c virus glycoproteins[J]. *J Gen Virol*, 1997, 78:2299-2306.
- [8] Ray B R, Meyer K, Ray R, *et al.* Suppression of apoptotic cell death by Hepatitis C virus core protein[J]. *Virology*, 1996, 226:176-182.
- [9] Rosa D, Campagnoli S, Moretto C, *et al.* A quantitative test to estimate neutralizing antibodies to the hepatitis C virus: Cytofluorimetric assessment of envelope glycoprotein 2 binding to target cells [J]. *proc Natl Acad Sci*, 1996, 93:1759-1763.