

HCV C 基因部分序列的克隆测序^①

吕 凌² 王 斌

(中山医科大学分子医学研究中心; 广州, 510089)

提 要 选取 5 株 HCV 共有序列合成一对 5' 端加端引物。上游加端为 *EcoR* I 切点, 下游加端为 *BamH* I 切点。以 RT-PCR 扩增一长为 578 bp 的 cDNA 作目的基因, 相应酶切后, 反向插入 pUC18 的多克隆位点, 构建 pUHC-C 重组体。分别或混合应用 HCV 特异内外引物和 pUC 特异通用引物以 PCR 扩增重组体, 扩增产物分子片段均同预期一致。测序表明克隆基因与 HCV-CHN 最同源, 在所测 C 区 497 个核苷酸及由之推导的 160 个氨基酸区域内, 其核苷酸与氨基酸同源性都是 97.5%。

主题词 克隆; 分子; 序列分析, DNA; 肝炎病毒组, 丙型; 聚合酶链反应

中图分类号 R 512.62; 523.1

丙型肝炎病毒(HCV)已被确认为肠道外非甲非乙型肝炎(PT-NANBH)的主要病原, 基因组为一正链极性的 RNA, 长 9 600 nt, 含一个开放阅读框架(ORF), 可编码 3 010 个氨基酸。依其编码的蛋白质不同, 该 ORF 被划分为若干个基因区, C 基因区是推定的核壳蛋白编码序列, 基因工程表达该序列的产物是目前 ELISA 检测抗 HCV 诊断试剂的主要成分^[1]。本文报道用 RT-PCR 自慢性丙型肝炎病人血中扩增该区段的大部分序列, 并进行定向分子克隆和序列测定, 以研究 HCV 基因组的变异。

1 材料与方 法

1.1 HCV 基因引物

根据 HCV-1^[2]、HCV-J^[3]、HCV-BK^[4]、HCV-TWN^[5]、HCV-CHN^[6]共有序列, 委托中国科学院上海细胞生物学研究所人工合成。C1/C2 用于扩增目的基因, C3/C4 用于鉴定克隆。为便于定向克隆, 在 C1 的 5' 末端改建了一个 *EcoR* I 切点, 在 C2 的 5' 末端改建了一个 *BamH* I 切点, 其序列和位置见表 1。

表 1 HCV 基因的引物序列

引物名称	核苷酸位置 ¹⁾	极 性	序列(5'-3')
C1	247-274	+	5'-GCGAATTC CCTTGTGGTACTGCCTGATA-3'
C2	798-825	-	5'-CAGGATCCCTGTTGCATAGTTCACGCCG-3'
C3	301-318	+	5'-TCTCGTAGACCGTGCACC-3'
C4	685-704	-	5'-CCGCATGTGAGGGTATCGAT-3'

1)核苷酸位置以毕胜利报道的中国河北株为准

1.2 pUC/M13 通用引物

Forward 为: 5'-GTTTTCCAGTCACGAC-3', Reverse 为: 5'-CAGGAAACAGCTATGAC-3', 均购自 Promega 公司。

1.3 RNA 提取

取一例男性、45 岁的输血后非甲非乙型肝炎病

人的新鲜血浆 500 μ L, 加 10 μ L SiO_2 悬液、1 mL 含胍裂解液, 室温下间歇摇匀 45 min, 待 SiO_2 自然沉降后, 小心吸除上层液相, 加 1 mL 80% 的乙醇悬浮洗涤, 10 000g 离心 3 min, 然后再吸除上清, 风干 10 min, SiO_2 为 Sigma 公司产品, 其悬液与含胍裂解液按文献^[7]配制。

① 广东省 920448.930386 号自然科学基金资助课题; ② 第一作者, 1958 年出生, 医学博士, 副研究员

1.4 逆转录

向干燥 SiO₂ 沉淀中直接加入 30 μL 逆转录混合液, 充分摇匀并保持悬浮, 43℃ 逆转录 1 h。逆转录混合液成分同前文^[9], 不同处是以 50 pmol 随机六聚寡核苷酸作为逆转录引物。

1.5 PCR 扩增

离心弃去逆转录液中的 SiO₂, 取 10 μL 上清加入 90 μL PCR 反应液进行 PCR 扩增, 条件同前文^[9], 不同处是以 C1 和 C2 为引物。

1.6 PCR 产物定向克隆

载体 pUC18 和 PCR 产物都以 *Bam*HI 和 *Eco*R I 双酶切, 经低熔点琼脂糖胶电泳回收, 加入 T₄DNA 连接酶进行连接。连接物转化钙化菌 *E. coli* JM109, 然后在含 IPTG/X-gal 的 LBA 平板上筛选白色菌落, 快速提取质粒 DNA 以限制性酶切和 PCR 鉴定插入序列。

1.7 DNA 序列测定

重组 HCV 序列采用双脱氧链末端终止法在 373A DNA 全自动测序仪 (ABI 公司) 上测定核酸序列, 测定时采用 Taq DNA 多聚酶及荧光标记的通用引物, 同一片段经正反两方向测定。

2 结果

C1 引物 5' 末端有 *Eco*R I 切点, C2 引物 5' 末端有 *Bam*HI 切点, 两者扩增产物长 578 bp, 用 *Eco*R I 和 *Bam*HI 双切后作为目的基因与同样酶切之 pUC18 载体连接, 构建重组质粒 pUHC-C。pUHC-C 长 3233 bp, *Bam*HI 单切后其电泳位置较线性化 pUHC18 空载体 (2686 bp) 后移, 并可被 *Eco*R I 和 *Pst*I 双切为两个片段, 一个长度为 586 bp, 略大于 578 bp 的目的基因, 另一个长度为 2647 bp, 略小于空载体 (2686 bp), 内切酶 *Cla*I 在目的基因第 440 nt 位置有一个切点而在载体上没有, 以此酶和 *Eco*R I 双切 pUHC-C, 可切出一个长度略大于 440 bp 的片段。内切酶 *Hind*III 在载体上有一个切点而在目的基因上没有, 以之单切 pUHC-C, 其结果与以 *Bam*HI 单切 pUHC-C 的结果一致, 电泳迁移位置相同 (图 1)。

在空载体 pUC18 的基因记位上, 通用引物 Forward 是 359-375 nt, Reverse 是 481-465。两引物配对扩增空载体, 产物长 123 bp; 扩增 pUHC-C, 产物长 670 bp, 后者长度约是 123 bp 产物与 578 bp 目的基因

之和。以 pUHC-C 为模板, 用 C1/C2 引物对扩增, 产物大小等于目的基因; 以 C3/C4 引物对扩增, 产物长 405 bp; 用 Reverse/C2 配对扩增, 产物长 602 bp; 用 Forward/C1 引物配对扩增, 产物长 646 bp, 用 Reverse/C4 引物配对扩增, 产物长 483 bp, 用 Forward/C3 配对扩增, 产物长 593 bp (图 2)。可见所有扩增的 PCR 产物其分子片段长度均与预期相符, 并且其克隆基因的插入方向也被证实为反向, 图 3 示 PCR 引物在 pUHC-C 中的记位及扩增时引物延伸的方向和扩增产物的长度。



图1 pUHC-C 酶切鉴定

S: *Spp*I/*Eco*R I; M: pBR322/*Msp*I; 1: 578 bp 目的基因; 2: pUC18/*Eco*R I + *Bam*HI; 3: pUHC-C/*Bam*HI; 4: pUHC-C/*Eco*R I + *Pst*I; 5: pUHC-C/*Eco*R I + *Cla*I; 6: pUHC-C/*Eco*R I



图2 pUHC-C 的 PCR 鉴定

1: C1 + C2, 578 bp; 2: Rev + C4, 483 bp; 3: For + Rev, 670 bp; 4: Rev + C2, 602 bp; 5: C3 + C4, 405 bp; 6: C2 + C3, 525 bp; 7: For + C3, 593 bp; 8: For + C1, 646 bp; 9: For + Rev/pUC18, 123 bp; M: Φ X174/*Hae*III。由上至下片段长度 bp 依次为: 1353, 1078, 872, 603, 310, 281, 271, 234, 194, 118, 72

将克隆的 HCV 基因进行序列测定 (图 4) 并与已发表的 HCV-J^[2]、HCV-J^[3]、HCV-BK^[4]、HCV-TWN^[5]、HCV-CHN^[6]、HC-J8^[10] 和 HC-J6^[11] 等分属 4 个型别的 7 株 HCV 序列进行同源性比较, 结果表明

本文 HCV 序列与毕胜利报道^[1]的最为同源。两者相比,在 ATG 前属于 5'-NCR 区的 43 个核苷酸(nt)完全一样,ATG 开始后的 479 nt(除去下游引物 28 nt)中有 12 个 nt 变异(见表 2),6 个为无义突变,6 个可导致由密码子推导的 4 个氨基酸发生改变,依次为第 107 位密码子由 AGT→AGA,导致丝氨酸→精氨酸,第 141 位密码子由 GTC→ATT,导致缬氨酸→异亮氨酸,第 142 位密码子由 GGC→CGG,导致甘氨酸→

精氨酸,第 157 位密码子由 GTT→GGT,导致缬氨酸→甘氨酸。此 479 nt 区域的核苷酸及所编码的 160 个氨基酸其变异率均为 2.5%。有义突变的位置都靠近高变的包膜蛋白编码区。

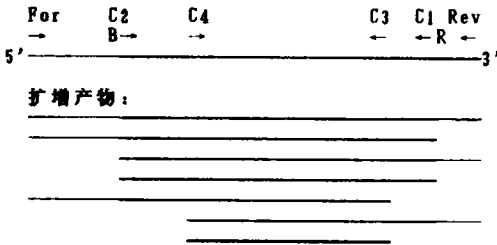


图3 PCR 引物标记及其扩增产物的长度

For=Forward 引物; Rev=Reverse 引物; R=*EcoRI*; B=*BamHI*; →:引物及其扩增方向; ---:载体 pUC18 的 DNA 序列; —:克隆 HCV 的基因序列

所用引物	长度
For+Rev	670 bp
For+C1	646 bp
C2+Rev	602 bp
C2+C1	578 bp
For+C3	593 bp
C4+Rev	483 bp
C4+C3	405 bp

表2 HCV C 基因部分序列的核苷酸变异情况

核苷酸位置 ¹⁾	变异情况	核苷酸位置 ¹⁾	变异情况	核苷酸位置 ¹⁾	变异情况
327	G→T	627	C→T	738	C→T
588	C→T	636	T→A	739	G→C
591	T→C	729	G→T	741	C→G
621	C→A	736	G→A	788	T→G

1)核苷酸位置以毕胜利报道的中国河北株为准

3 讨论

HCV 是第一个以反向生物学方法阐明的人类病毒,其病毒颗粒迄今未被公认但基因序列却通过分子克隆而被阐明^[2]。HCV 基因组长 9 600 nt,携一个 ORF,可分为若干基因区^[2]。其中 C 区编码 HCV 的核壳蛋白,是丙肝 ELISA 诊断试剂的一种主要抗原^[1]。

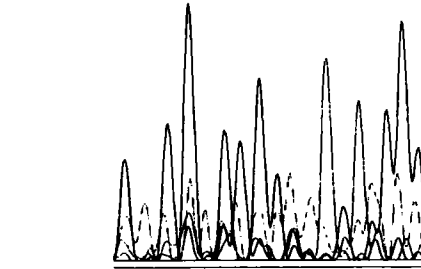


图4 pUHC-C 测序及同 HCV 代表株比较

---:相同; - - -:中间的核苷酸缺乏

GAATGAGCGGAATGTACCC 换成互补链;

723 733 741 nt

GGGTACATTCGGCTCATTC pUHC-C

-----G-----G-CG	HCV-CHN ^[6]
-----G-----G-CG	HCV-TWN ^[5]
-----TG-CG	HCV-J ^[3]
-----G-CG	HCV-BK ^[4]
-----A-----G-CG	HCV-US ^[2]
-----CATCC-T-----G--G	HCV-J8 ^[10]
-----C--TG--G-AG	HCV-J6 ^[11]

单股 RNA 的 HCV 基因在人体内含量极微,故采用 RT-PCR 扩增 HCV-cDNA 作为目的基因。扩增片段 5' 端起始于距 ATG 上游的 71 nt 处,3' 端止位编码第 170 位氨基酸的密码子位置^[6]。一般认为 HCV 多蛋白的头 190 个氨基酸是 HCV 核壳蛋白^[9],故上述扩增序列包括了其头端 89% (170/190)。

以加端 PCR 引物扩增目的基因,C1 引物 5' 端含 *EcoRI* 切点,C2 引物 5' 端含 *BamHI* 切点,所得目的基因带两个不同粘性末端,因此能定向克隆,即在

pUC 18的 *EcoRI* 和 *BamHI* 位点中反向插入。

用 PCR 技术鉴定了克隆 HCV 基因的长短、特异性及插入方向。对 pUHC-C 以 C1和 C2引物扩增,产物大小等于目的基因,以内引物 C3和 C4扩增,产物为404 bp,以通用引物扩增,产物为670 bp(空对照为123 bp),上述三种产物的长度大小均同预期相符,故克隆基因长短和特异性得以鉴定。以载体特异的通用引物和 HCV 特异引物配对,结果延伸方向相对者有扩增,相背者无扩增,因此克隆基因插入方向也得以鉴定。

以限制性内切酶鉴定了重组质粒上的酶切点,结果与实际情况一样。目的基因上有单个 *ClaI* 切点,载体上有单个 *Hind III* 和 *PstI* 切点,目的基因与载体连接处有单个 *BamHI* 和 *EcoRI* 切点。单切或双切的结果证实了这些位点的存在。

DNA 测序表明,本文 HCV 基因与文^[6]报道的序列最为同源,同属于第 I 型 HCV,在所测 C 区 497 nt 及由之推导的160个氨基酸区域内,其核苷酸与氨基酸同源性都是97.5%。至于变异的4个氨基酸对 HCV 核壳蛋白抗原性影响如何还有待研究。

参 考 文 献

- 1 Chiba J, Ohba H, Matsuura Y, *et al.* Serodiagnosis of hepatitis C virus infection with an HCV core protein molecularly expressed by a recombinant baculovirus. *Proc Natl Acad Sci USA*, 1991, 88(11):4641
- 2 Houghton M, Weiner A, Han J, *et al.* Molecular biology of the hepatitis C viruses: implications for diagnosis, development and control of viral diseases. *Hepatology*, 1991, 14(2):381
- 3 Kato N, Hijikata M, Ootsuyama Y, *et al.* Molecular cloning of the human hepatitis C virus genome from Japanese patients with non-A, non-B hepatitis. *Proc Natl Acad Sci USA*, 1990, 87

(24):9524

- 4 Takamizawa A, Mori C, Fuke I, *et al.* Structure and organization of the hepatitis C virus genome isolated from human carriers. *J Virol*, 1991, 65(3):1105
- 5 Chen PJ, Lin MH, Tai KF, *et al.* The Taiwanese hepatitis C virus genome: sequence determination and mapping the 5' termini of viral genomic and antigenomic RNA. *Virology*, 1992, 188(1):102
- 6 毕胜利, 白宪鹤, 刘崇柏, 等. 我国丙型肝炎病毒河北株基因组的克隆及 cDNA 全序列的测序与分析. *中华实验和临床病毒学杂志*, 1992, 6(4):425
- 7 Boom, R, Sol CJA, Salimans MMM, *et al.* Rapid and simple method for purification of nucleic acid. *J Clin Microbiol*, 1990, 28(3):495
- 8 吕 凌, 姚集鲁, 彭文伟, 等. 非甲非乙型肝炎121例血清丙型肝炎病毒抗体和血浆丙型肝炎病毒 RNA 的研究. *中华传染病杂志*, 1993, 11(4):187
- 9 金冬雁. 丙型肝炎病毒分子生物学研究动态和发展趋向. 见: 侯云德, 金冬雁主编. *现代分子病毒学选论*. 北京: 科学出版社, 1994. 244~260
- 10 Okamoto H, Kurai K, Okada S, *et al.* Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates; comparative study of four distinct genotypes. *Virology*, 1992, 188(1):331
- 11 Okamoto H, Okada S, Sugiyama Y, *et al.* Nucleotide sequence of the genomic RNA of hepatitis C virus isolated from a human carrier; comparison with reported isolates for conserved and divergent regions. *J Gen Virol*, 1991, 72(pt 11):2697

(1996-03-01收稿 1996-08-28修回)

CLONING AND SEQUENCING OF THE PARTIAL C SEQUENCE FROM HCV

Lu Ling Wang Bin

(The Research Center of Molecular Medicine, Sun Yat-sen University of Medical Sciences, Guangzhou 510089)

From published data of five strain of HCV, two conservative sequences were applied to synthesize a pair of primers, on the 5'-termini of which, a restrictive site was constructed. Position of upstream primer in HCV genome is 247~274, and a *EcoR* I site was added. Position of downstream is 825 R-789 R, and a *BamH* I site added. By RT-PCR, a cDNA fragment of 578 bp was amplified from plasma that is positive to anti-HCV. By directed cloning, the fragment was inserted reversely into the polycloning site of the vector pUC18, and a recombinant plasmid pUHC-C was generated. For identification of the insert, PCR has been used employing two nested primer pairs specific to HCV, one primer pair specific to pUC18, and 4 mixed primer pairs from both. In each reaction, an anticipated product was yielded that proves the insert to be the HCV C gene and its direction is reverse. This fact was also supported by the result of the restrictive site analysis. By sequencing the genotype of HCV was identified to be HCV- I, herein 12 nt mutation exist, of which 6 may affect their coding amino acids. 97.5% is the homology of the cloned cDNA and its deduced peptide sequence to HCV-CHN.

Subject headings molecular cloning and DNA sequencing/methods; polymerase chain reaction/methods; hepatitis C virus; DNA

~~~~~

·新成果·

## 聚 $\alpha$ -氨基酸缓释药物系统的研究

课题负责人 潘仕荣

由潘仕荣教授等人完成的聚  $\alpha$ -氨基酸缓释药物系统的研究,成功地合成了聚  $\alpha$ -氨基酸,即白氨酸-谷氨酸甲酯-谷氨酸共聚物(甲)和白氨酸-谷氨酸的苄酯共聚物(乙)并研究了它们的生物降解性,其生物降解性取决于氨基酸的组成和比例,用共聚物甲制备了缓释药囊、缓释药片,用共聚物乙制备了缓释微球,研究其体外释药规律。找出了氨基酸的亲水性,释药速率,药物溶解度与时间等的规律,其中有些规律建立了方程式。本项目对促进国内缓释药物研究理论的发展具有重要意义,制备的缓释药物对制药产生的发展具有很大的社会效益和潜在的经济效益。该成果1994年获国家教委科技进步三等奖。

(陈丽芳)