

23T-242

6

R373.21

戊型肝炎病毒(HEV)开读框架2中1.3 kb cDNA 的克隆与序列分析

蒋琳¹, 杨恭², 沈心亮¹

¹(卫生部兰州生物制品研究所, 兰州 730046)

²(兰州大学生物系, 兰州 730000)

摘要:以戊型肝炎(Hepatitis E)病人高热期血清为原材料,设计特定引物,经 RT-PCR 扩增获得开放阅读框 2(ORF2)1.3 kb 基因片段,用 *EcoR* I 和 *Bam* H I 插入克隆载体 pUC18,测定了该克隆基因片段的核苷酸序列并进行了比较分析。其片段长度为 1 309 bp,其 A = 249(19.2%), G = 318(24.29%), T = 339(25.9%), C = 403(30.79%);与印度株和缅甸株的同源性在 92% 以上,而与乍得和墨西哥株的同源性分别为 90.7% 和 77.4%。

关键词:戊型肝炎病毒; cDNA 克隆; 测序

中图分类号: R512.36 **文献标识码:** A **文章编号:** 1003-5125(2000)03-0237-06

戊型肝炎病毒(HEV)在形态上属于杯状病毒科^[1,2]。为粪口传播的肝炎病毒之一。1983年,由苏联学者用急性期病人粪便口服液进行志愿者和动物实验^[3,4],发现粪便中病毒颗粒能与志愿者恢复期血清发生凝集。90年代初,HEV ET1.1 克隆及序列分析^[5],推动了 HEV 分子生物学研究的进展。来自缅甸^[5,6]、中国^[7]、墨西哥^[8]、乍得^[9]、印度^[10,11]等地区的 HEV 毒株经部分或全部基因序列分析表明,不同地区分离的 HEV 的基因结构相似,均为单股正链 RNA,由 7 194 个核苷酸和 3'端的 poly(A)尾巴组成。整个基因组有三个开放性阅读框架(ORFs):ORF1 为 HEV 的非结构基因编码区;ORF2 是 HEV 的主要结构基因编码区,起于核苷酸的 5 147 位,延伸 1 980 bp,止于 poly(A)上游的 631 个核苷酸;ORF3 只含有 369 bp,与 ORF1 轻微重叠(1 bp),而与 ORF2 广泛重叠(328 bp),其编码产物与该病毒的特异性免疫反应有关。

由于戊型肝炎病毒到目前为止尚没能组织培养成功,病毒稳定性比其它种类 RNA 病毒差,基因组的克隆相对难度大。目前全世界克隆成功的戊型肝炎病毒全基因组大多为灵长类动物感染穿刺后克隆成功。本研究以人类标本(戊型肝炎病人血清)为原材料,提取病毒 RNA;根据 HEV 基因组 ORF1, ORF2, ORF3 的特点选择编码产物与病毒的特异性免疫反应有关的 ORF2,设计合成特定引物,通过 RT-PCR 获得目的基因扩增;将扩增的基因克隆于质粒载体 pUC18 中,进行了序列测定和分析。本实验的完成为下一步对该基因片段进行真核与原核细胞的表达以及研制戊型肝炎基因工程疫苗奠定了基础。

收稿日期:1999-06-28,修回日期:1999-10-13

作者简介:蒋琳(1962年-),女,上海人,研究员,研究方向:分子生物学。

1 材料与方法

1.1 菌种与质粒

大肠杆菌 JM109, 来自中国预防医学科学院病毒研究所, pUC18 购自华美生物工程公司。

1.2 试剂与临床标本

MMV 反转录酶, EDTA 为 Gibco BRL 产品; *EcoR* I, *Bam* H I, Taq DNA 聚合酶, T4 DNA 连接酶和 dNTP 购自 Promega 公司; DTT RNase, RNasin 购自 Sigma 公司; 酵母提取物 (yeast extract) 和蛋白胨 (tryptone) 为 Unipath Ltd (England) 产品; 琼脂糖 (agarose)、低熔点琼脂糖 (LMP-agarose)、Trizol 试剂购自德国宝灵曼公司; 其它常规试剂均为国产分析纯。HEV 高热期病人血清来自新疆维吾尔自治区卫生防疫站。

1.3 引物设计

上游引物 S3: 5'-CGTCCGTTGACATGAATTCA-3' (20 bases, 下划线为 *EcoR* I 酶切位点), 下游引物 AS3: 5'-CGGGATCCCGATAACTCCCGAGTTTTTA-3' (27 bases, 下划线为 *Bam* H I 酶切位点), 由中国科学院微生物研究所合成。

1.4 DNA 常规操作和质粒转化

见文献^[12]和^[13]。

1.5 病毒 RNA 的提取

按文献^[3]报道的方法进行。

1.6 病毒 RNA 的 RT-PCR

在提取的病毒 RNA 中加入 HEV 特异引物及反转录酶 MMV 在 65 °C 加热 5 min, 然后 42 °C 反转录 40 min。PCR 循环为: 94 °C 3 min, 45 °C 1 min, 72 °C 2 min, 一个循环; 94 °C 40 s, 45 °C 50 s, 72 °C 1 min 30 s, 35 个循环; 72 °C 延伸 10 min。

1.7 重组质粒的构建与克隆基因 DNA 序列的测定

用低熔点琼脂糖回收 PCR 扩增片段, *EcoR* I 和 *Bam* H I 分别酶切目的片段和质粒 pUC18, T4 DNA 连接酶连接, 转化于感受态大肠杆菌 JM109, 根据蓝白斑筛选重组克隆, *EcoR* I 和 *Bam* H I 酶切及 PCR 扩增鉴定重组质粒。

碱法提取重组质粒, PEG 纯化, 用双脱氧末端终止法 (荧光标记, Model 373A 自动测序) 测定克隆基因的 DNA 序列。

2 结果与讨论

2.1 RT-PCR 与重组质粒 (pHEVORF2-1.3) 的酶切鉴定 (图 1)。



图 1 PCR 产物与重组质粒 pHEVORF2-1.3 的限制酶切电泳图

Fig. 1 PCR product and restriction endonuclease analysis of recombinant pHEVORF2-1.3

Lane 1: DNA/EcoRI & Hind III Marker. Lane 2: PCR product. Lane 3: pUC18/EcoRI. Lane 4: recombinant plasmid of pHEVORF2-1.3/EcoRI & BamHI. Lane 5: pHEVORF 2-1.3

本实验中扩增并克隆的 HEV 1.3 kb cDNA 为其 ORF2 下游部分 5 818-7 126 部分。该段基因主要编码戊肝病毒外壳蛋白, 其表达的产物与 HEV 保护性抗原密切相关。

2.2 克隆基因核苷酸序列测定及序列比较分析

序列比较图见图 2。


```

730      740      750      760      770      780      790      800      810      820      830
CIBATGATGTGCTTGGCTCTCTCACGGCTGCGGAGTATGACCACTTACGGCTCTTCGACGGGCGCCAGTCTATGTTCTGACTCTGTGACCTTGGTT
.....G.....T.....T.....T.....T.....T.....T.....T.....T.....C.....C.....
.....T.....T.....T.....T.....T.....T.....T.....T.....T.....C.....C.....
.....T.....C.....C.....T.....C.....C.....C.....C.....C.....C.....
A.G...TAC.TTGGCT.TC.CTCACTG.A..C.AGT.TG.OC.GTCCA.TTACGGGTCG..AA.TGG.CCGGTTTATATCTGGACAG.ETGACTTTGG..AA.
.....T.....T.....T.....T.....T.....T.....T.....T.....T.....C.....
.....T.....T.....T.....T.....T.....T.....T.....T.....T.....C.....

840      850      860      870      880      890      900      910      920      930
AATGTTGCGACCGCGCGGCGGCGGCTTCCCGGCTCGCTGGACTGGACCAAGGTCACACTTGTATGGTGGGCGCCCTCTCCACCATCCAGCAGTATTGGAAGACCTT
.....A.....T.....C.....CC.....
.....T.....T.....C.....T.....T.....A.....
.....G.....T.....C.....T.....A.....T.....T.....T.....T.....
.....C.....T.....T.....C.....C.....C.....T.....T.....T.....
GT..CGA.TGG..CGCA.GCC.TA.CCGGATC.CTGA.TGG.CC.AAGTCAC.CTCGAC.GGC.G.C..T..CGA.TGTGGAG..AT.T.CCAA..CATT...
.....T.....T.....C.....C.....
.....A.....T.....C.....T.....C.....A.....

940      950      960      970      980      990      1000     1010     1020     1030     1040
CTTGTCTCGCGCTCCGGGTAAAGCTCTCTCTGGGAGGCGAGGTACANCTAAAGCGGGTACCCCTTATAATTATAACACCACTCTAGCGACCAACTGCTTG
.....T.....T.....T.....T.....T.....T.....T.....T.....T.....T.....C.....
.....T.....C.....C.....G.....C.....T.....T.....T.....A.....
.....C.....T.....C.....G.....T.....C.....T.....T.....CA.....
.....G.....T.....C.....GGC.....C.....A.....
TG.GC...CC.TT.GTG..AAGCTCTC..TT.GGGA..CC.GGCA...AAGC..GTATCTTA.A..T.A...CT..TG..AG.GA.C.GATT..A...
.....T.....C.....T.....T.....T.....T.....T.....T.....T.....T.....C.....
.....T.....T.....T.....T.....T.....T.....T.....T.....T.....C.....

1050     1060     1070     1080     1090     1100     1110     1120     1130     1140
TGGAGATGCCCGCGCGCATCGGGTTGCTATTTCCACTTTCACCACTAGCCCTGGGTGCTGGCCCCGTCTCCATTTCGCGGTGCTGTTTAGCCCCCACTCT
.....T.....C.....C.....T.....T.....C.....
.....T.....C.....T.....T.....TG...C.....G.....
.....C.....T.....G.....C.....T.....T.....A.....A.....T.....C.....C.....C.....
AAA.TGC...G.CAT.GGCTC.CCAT.TCAA..TA.AC...GGCTTGG..CC.S.CGGT..C.ATTC.G.G..C.OE.T.T.GGCTC.A.G.TC.G.C
.....C.....C.....T.....T.....T.....T.....C.....C.....
.....T.....T.....T.....T.....T.....C.....

1150     1160     1170     1180     1190     1200     1210     1220     1230     1240     12
GGGCTAGCATTGTGAGGATACCTTGGACTACCTGCCGCGGCCATACTTTTGATGACTTCTGCCAGAGTGGCGCCCCCTGGGCTCCAGGGCTGTGCTTT
.....A.....C.....G.....
.....CC.....A.....C.....T.....T.....T.....C.....
.....TT.....C.....T.....T.....T.....T.....G.....
CT..GCTGTC..GAG..TAC.TTTGATT.TCGGGGCGGGGCA.ACATT.GA...CTT..G.CCTGA.TGCC..GCTTTAGGCT..AGGTT.TGC.TTCCA
.....A.....C.....C.....T.....T.....G.....

50      1260     1270     1280     1290     1300     1310     1320     1330     1340     1350
CCAGTCTACTGTCCGCTGAGCTTCAGGGCTTAAGATGAAGGTTGGTAAAACCTCGGGAGTTA*AGTTTATTTGCTTGTGCCCCCCCTCTTTCTGTGTGCTTATTTTC
T.....C.....C.....C.....C.....C.....C.....C.....C.....C.....C.....C.....C.....
T.....C.....G.....T.....T.....T.....C.....C.....C.....
GTCAA..GTC.CT.AGCTC.AG.GC.TTAAAGTT.A.GT..GTAAG.CTCGGA.TT..AG.TTA..TGGCTG.GCCCA..TA..A.A...C..A..TCC.T
.....G.....
T.....T.....T.....T.....T.....T.....T.....T.....T.....T.....T.....T.....T.....T.....T.....T.....T.....T.....T.....

1360     1370     1380     1390
TCATTTCTGGGTTCTGGGCTCCCTGAAAAAATAAAAAAAAAA
Hevsq00
.....T.....10-1-1
.....T.....14-1-1
.....T.....1-1-1
.....AT...ITTECT..GT.C.G.GCTCCCTG.....19-1-1
.....15-1-1
.....16-1-1
.....2-1-1

```

图2 克隆基因核苷酸序列测定结果及序列比较

Fig.2 Cloned cDNA sequence and comparison among different isolates

克隆基因的序列测定结果为:序列长 1 309 bp, A = 249 (19.2%), T = 339 (25.9%), C = 403 (30.79%), (A + T) = 44.92%, (G + C) = 55.08%。序列(Hevseq00.seq)比较分析表明:与印度 1(10-1-1.seq, 基因 bank 号:U22532, 下同)和印度 2(14-1-1.seq, X98292)的同源性分别为 92% 和 93.3%;与缅甸 1(1-1-1.seq, M73218)和缅甸 2(19-1-1.seq, D10330)的同源性分别为 92.8% 和 92.6%;与乍得株(16-11.seq, U62121)的同源性为 90.7%;与墨西哥株(15-1-1.seq, M74506)的同源性为 77.4%;而与另一中国株(2-1-1.seq, L08816)的同源性为 97.9%。说明作为结构蛋白的 ORF2 编码区在亚洲株之间是相对稳定的,而与非洲株之间的差异则较大。不同 HEV 毒株的 ORF2 核苷酸序列变化决定了它们编码蛋白的抗原决定簇的差异,而这种差异直接影响到戊型肝炎的血清学诊断和疫苗的研制。

参 考 文 献

- [1] Bradley D W. Hepatitis E: Epidemiology, aetiology and molecular biology [J]. *Rev Med Virol*, 1992, 2:19-28
- [2] Taarve S A, Emerson S U, Reyes G R, *et al*. Characterization of a prototype strain of Hepatitis E virus [J]. *Proc Natl Acad Sci*, 1992, 89:559-563
- [3] Favorov M O, Fields H A, Purdy M M, *et al*. Serologic identification of Hepatitis E virus infection in epidemic and endemic setting [J]. *J Med Virol*, 1992, 36:246-249
- [4] Chauhan A, Jmeel S, Dilawari J B, *et al*. Hepatitis E virus Transmission to volunteer [J]. *Lancet*, 1993, 341:149-150
- [5] Tam A W, Smith M M, Egueraetal M. Hepatitis E virus: Molecular cloning and sequencing of the full-length viral genome [J]. *Virology*, 1991, 185:120-131
- [6] Aye T T, Uchina T, Ma X, *et al*. Sequence and structure of hepatitis E virus isolated from Myanmar [J]. *Virus Genes*, 1993, 7:95-109
- [7] Aye T T, Uchina T, Ma X, *et al*. Complete Sequence of hepatitis E virus isolated from the Xinjiang epidemic (1986-1988) of China [J]. *Nuc Acids Res*, 1992, 20:3512
- [8] Huang CC, Nguyen D, Fernander J, *et al*. Molecular cloning and sequencing of the Mexico isolate of Hepatitis E virus [J]. *Virology*, 1992, 191:550-558
- [9] Van Cuyck-Gandre H, Zhang H Y J, *et al*. Characterization of Hepatitis E virus from Algeria and Chad by partial genome sequence [J]. *J med Viro*, 1997, 53:340-347
- [10] Panda S K, Nanda S K, Zafrullah M, *et al*. An Indian strain of hepatitis virus: cloning, sequence, and expression of structural region and antibody responses in sera from an area of high-level HEV endemicity [J]. *J. Clin. Microbiol*, 1995, 33: 2653-2659
- [11] Donati M C, Fagan E A, Harrison T J, in: Rizzetto M, Purcell R. H, Gerin J L, Verme G (eds.); *Viral Hepatitis and Liver Disease* [M], 1997, 313-316
- [12] Sambrook J, Fritsch E F, Maniatis T. *Molecular cloning: A laboratory manual* 2nd ed. Cold Spring Harbor Lab Press, New York, 1989
- [13] Ausubel F M, Brent R, Kingston R E, *et al*. *Short Protocols in Molecular Biology*. (1995, Third Edition)

Cloning and Sequence Analysis of the 1.3 kb cDNA Fragment from Open Reading Frame 2 (ORF2) of Hepatitis E Virus (HEV)

JIANG Lin¹, YANG Gong², SHEN Xin-liang¹

¹(Lanzhou Institute of Biological Products, Lanzhou 730046, China)

²(Biological Department of Lanzhou University, Lanzhou 730000, China)

Abstract: The serum from patients infected with Hepatitis E was prepared, the RNA of Hepatitis E virus was purified. The interested cDNA fragment (1.3 kb) was obtained by RT-PCR, cloned in to the plasmid pUC18 and sequenced by the dideoxynucleotide method (fluorescein labeled dNTP). The result of sequencing suggested that the cDNA fragment was 1 309 bp, its adenine, guanine, thymine and cytosine were 249 (19.2%), 318(24.29%), 339(25.9%) and 403(30.79%), respectively. The nucleic acid sequences of the cDNA were 92% to 93.3% homologous to the Indian and Burmese strains, 90.7% and 77.4% homologous to Chad and Mexican strains. The ORF2 of HEV is relatively stable between the Asian strains.

Key word: Hepatitis E virus; cDNA cloning; Sequencing

* * The sequences have been submitted to Genbank: AF141652 (login NO.262766)