维普资讯 http://www.cqvip.com

第16卷1期 中国病毒学 16(1):28-33
2001年3月 VIROLOGICA SINICA Mar 2001

新型戊肝病毒ORF3和部分ORF1基因的序列分析*

王佑春[1],张华远[1],辜文杰[1],Roger Ling[2],李河民[1],Tim J Harrison[2]

(1.中国药品生物制品检定所,北京 100050;
2.Center for Hepatology, Royal Free Campus, University College London, Hampstead, London NW3 2PF, UK)

Sequence Analysis of ORF3 and Partial ORF1 Region from Two Patients Infected with New Genotype of Hepatitis E Virus(HEV)

WANG You-chun[1], ZHANG Hua-yuan[1], GU Wen-jie[1], Roger Ling[2], LI He-min[1], Tim J Harrison[2]

(1. *National Institute for the Control of Pharmaceutical and Biological Products, Beijing* 100050, *China*;
2 *Center for Hepatology, Royal Free Campus, University College London, Hampstead, London NW3 2PF, UK*)

Abstract: ORF3 and partial ORF1 regions were amplified with RT-PCR from two patients (T1 and T11) infected with new genotype of hepatitis E Virus. The PCR products were cloned and sequenced. The results showed that G-C rich region in ORF3 was deleted when amplified with normal PCR reaction. However, PCR reaction containing G-C melt solution can overcome this problem. The sequence analysis showed that T1 and T11 belong to a new genotype of HEV which differs from genotype I, II and III reported. T1 and T11 have 79%~82%, 80%~81% and 83%~85% identical to genotype I, II and III respectively.

Key words: Hepatitis E virus; Genotype; Homology; Open Reading Frame

摘要:用RT-PCR方法对2例感染新型HEV的样品进行了检测,并对PCR产物进行了克隆及测序。检测结果显示用常规PCR检测易造成ORF3中GC丰富区的缺失,但在常规PCR反应液中加入GC melt溶液可成功地扩增GC丰富区。序列分析显示T1和T11属于同一基因型,但不同于已报道的HEV I型、II型和III型,为一新的基因型。T1和T11与I型在该区的核苷酸同源性为79%~82%;与II型的同源性为80%~81%;和III型的同源性为83%~85%。

关键词:戊型肝炎病毒;基因型;同源性;开放读码框架
中图分类号:R373.2 **文献表识码**:A **文章编号**:1003-5125(2001)01-0028-06

戊型肝炎病毒(HEV)是肠道传播性非甲非乙型肝炎的主要病因,主要经污染的水源传播,也可通过饮食传播。该病毒即能引起大规模的爆发流行,也可引起急性散发型病毒型肝炎,国际上已对来自于不同地区的十几株戊肝病毒进行了全基因的克隆及测序,序列分析显示目前戊肝病毒可分为三个基因型,并有明显的区域性。I型来自于亚洲和非洲;II型来自于墨西哥;III型来自于美国[1]。最近我们自散发性戊肝病人分离出IV型HEV[2],并对IV型HEV的ORF2部分基因序列进行了克隆及序列分析,本研究对两株IV型HEV的ORF3区进行克隆及序列分析,以研究该区基因及氨基酸的变异程度。

1 材料与方法

1.1 粪便样品

样品T1采自于北京地区一位33岁的男性急

收稿日期:1999-12-6,修回日期:2000-11-27
* 基金项目:部分由英国Wellcome Trust基金组织资助
作者简介:王佑春(1964-),男,山东茌平县藕,研究员,医学和哲学博士,主要对乙型、戊型、庚型肝炎病毒及TTV进行分子生物学的研究。

维普资讯 http://www.cqvip.com

性肝炎病人;样品T11采自于北京地区一位39岁的男性急性肝炎病人。通过对样品T1和T11 HEV ORF2基因片段的检测,证明T1和T11为感染新型HEV的病人[2,3],因此决定对其ORF3片段进行扩增及序列分析。

1.2 引物设计

按照新型HEV的基因序列,设计了一套巢式PCR引物,该引物扩增片段的预计大小为533 bp。

外引物为:BO13: 5'-TTGTATGCTGGTGTGGTTGTAG-3'
EO13: 5'-TTGTATGCTGGTGTGGTTGTAG -3';
内引物为:BI13: 5'-GGATGTTGTTTCGCAGGTTTATG-3'
EI13: 5'-AGCAGGAGCCACA GCAGTCA-3'.

1.3 粪便样品处理及病毒RNA提取

取5 g粪便样品加PBS 50 mL,并彻底震荡使之悬浮,然后4℃条件下3 000 r/min离心10分钟,取上清并保存在-70℃冰箱中备用。用QIA病毒RAN提取试剂盒按操作说明提取病毒RNA,140 μL粪便悬液可获得50 μL纯化的病毒RNA液体。

1.4 反转录

9 μL纯化RNA与25 pmol PCR的反义外引物(EO13)混合。在70℃中加热10 min,然后冰浴2 min。再离心后加入4 μL 5·AMV缓冲液,1 μL 10 mmol/L dNTP,0.75 μL RNasin(40 u/μL),0.2 μL 100 mmol/L DTT,然后用DEPC水调整体积至20 μL,并在37℃的条件下反应1 h。

1.5 PCR反应

1.5.1 常规PCR反应 第一轮PCR反应液为10 μL反转录液,二个外引物各25pmol,5 μL 10· PCR反应液,4 μL 25 mmol/L $MgCl_2$,及5u Tag酶,并用无菌水调整终体积至50 μL,然后加2滴石蜡油,并在PCR扩增仪进行PCR反应,扩增参数为94℃ 1 min,50℃ 1 min,72℃ 2 min,共35个循环。取第一轮PCR反应产物5 μL,25 μL内引物及其它与第一轮相同的反应液进行第二轮PCR反应,具体反应同第一轮。

1.5.2 G-C丰富区的PCR扩增:在常规PCR反应中加入10 μL的5 mol/L GC melt溶液(Invitrogen公司)和5 μL的二甲基亚枫,其他反应液和扩增参数同常规PCR反应。

1.6 克隆及测序

PCR产物用2%的琼脂糖胶电泳检测,将预计大小相当的条带切下,并进行纯化,然后连接到pCR[II] T载体上,具体的克隆筛选及测序方法见参考文献[4]。每一个产物均筛选2-3个克隆进行双向测序,而且对部分区域进行手工测序核实。

1.7 序列分析

用Clustal W程序的基因序列进行对准比较,并用PHYLIP软件进行基因进化树的分析。

2 结果

2.1 常规PCR检测结果

用常规PCR对样品T1和T11检测,其预计PCR产物的大小为533 bp,但实际PCR产物仅为460 bp左右,对该产物进行克隆测序发现该产物为HEV基因序列(见图1中FT1和FT11序列),与其他序列相比缺失71个核苷酸,在缺失的71个核苷酸中GC含量为81.7%(58/71),为GC丰富区,用常规PCR法较难扩增。

2.2 G-C丰富区的PCR扩增

用该法对样品T1和T11均进行了成功的扩增,并对扩增产物进行了克隆和测序(见图1中T1和T11序列)。由表1可见T1和T11与其它序列相比在141位均插入一个T,并对该部分区域进行手工测序核实。

2.3 序列分析

将T1和T11序列与已报道的11株全基因序列进行了比较,B1(M73218),B2(D10330),P1(M80581),C1(L25547),C2(M94177),C3(D11093),I1(X98292)和I2(X99441)为I型;M1(M74506)为II型;U1(AF060668)和U2(AF060669)为III型。T1和I型的在该区的核苷酸同源性为79%~80%;和II型的同源性为80%;和III型的同源性为83%~84%;T11和I型的同源性为81%~82%;和II型的同源性为81%;和III型的同源性为85%。T1和T11在该区的核苷酸的同源性为90%,I型之间各基因序列在该区的同源性在95%以上(见表1)。

2.4 基因进化树分析

将T1和T11序列与已报道的11株全基因序列进行基因进化距离计算并显示其基因进化树(图2),由图2可见B1,B2,P1,C1,C2,C3,I1和I2形成一个分支;M1单独为一分支;U1和U2形成一个分支;T1和T11形成另外一个分支,不同于已报道的3个基因型,为一新的基因型。

维普资讯 http://www.cqvip.com


```
                    *        20         *        40         *        60
B1    : ..............A...............................G...............
B2    : ................T.................T.........G...............
C1    : ..............................................G...............
C2    : .............A...............................G..........C....
C3    : ....C.........................................G.............T.
I1    : .......T.......................................................
I2    : ....C.....C...................................G...............
P1    : ..............................................G...............
M1    : ..........G..T..G..............A........C......A...G......T.
U1    : .A..CAG...C........A........T...........G.....CA.C..C.......
U2    : .A...AG...C.....G..A........T...........G.....CA............
FT1   : .C...AG......TT.G..A..............G.....T.....CA....C.......
FT11  : .T...AG......TT.G..A...........A..G.....T......A....A.......
T1    : .C...AG......TT.G..A..............G.....T.....CA....C.......
T11   : .T...AG......TT.G..A...........A..G.....T......A....A.......
        GGGTTTCCCCTGGGCTCGTTCATAACCTGATTGGCATGCTACAGACTGTTGCTGATGGCA

                    *       100         *       120         *       140
B1    : ..A..A........C...T..............T...........................
B2    : ..A..A..G.....C...T..............T...........................
C1    : ..G...........C..................C..A........................
C2    : ..G...........C..................C...........................
C3    : ..G...........C...T..............C...........................
I1    : ..A...........C..................C...........................
I2    : ..A..A........C..................C...........................
P1    : ..G...............................TC..........................
M1    : .........A.A........T...C.C.....T...CAC...TCT...............
U1    : A....A.....A.....T..C........C.....ACAG.....................
U2    : A....A..............T........C.....ACAA.....................
FT1   : A..........C........T..T.GC..C..T..A.AC.....................
FT11  : A..........C........T..C.....T..A..A.A......................
T1    : A..........C........T..T.GC..C..T..A.AC.....................
T11   : A..........C........T..C.....T..A..A.A......................
        GTTAAGCCWGTGCTTGACCTGACAAATTCAATCATGTGTCGGGTGGAATGAATAACATGT

                    *       180         *       200         *       220
B1    : ...............................C..............T...........
B2    : ..............A..................C..............T........T..
C1    : .C.............................C..............................
C2    : ..............................C..............................
C3    : ..G............................C.........................CG..
I1    : ...........................C...................T........T..
I2    : ...............................C..............T...........
P1    : ..............................C..............................
M1    : .....C............A.....C.....C.....T.....T..............T..
U1    : ---ATC............A..G..G..C........T.......................
U2    : ---ATC............A..G..G..C...........T..G..C..T...........
FT1   : GAT..C.........T.....G..C..C....TC.....T..G..C...........T.-
FT11  : GAT..C.........T.....G..C..C....T...T.....T..C..T...........
T1    : GAT..C.........T.....G..C..C....TC.....T..G..C...........T..
T11   : GAT..C.........T.....G..C..C....T...T.....T..C..T...........
        TTCGCGACCATGCGCCCTCGGCCTATTTTGTTGTTGCTCCTCATGTTTCTGCCTATGCTG

                    *       260         *       280         *       300
B1    : ..............................................................
B2    : ..............................................................
C1    : ..............................................................
C2    : ...............C..............................................
C3    : ..............................................................
I1    : ..............................................................
I2    : ..............................................................
P1    : ..............................................................
M1    : ...................................A.....................
U1    : ............T..C.....................G.................A....
U2    : ...................................G.....................
FT1   : ----------------------------------------------.-----.........
FT11  : ----------------------------------------------------.........
T1    : ............T..C..C....A----......-...G.-....................
T11   : ............T.......................G.....................
        GCCGTCTGGCCGCCGTCGTGGGCGGCGCAGCGGCGGTTCCGGCGGTGGTTTCTGGGGTGA
```

维普资讯 http://www.cqvip.com


```
                  *        340         *        360         *        380
B1    : ..................................................................
B2    : ..................................................................
C1    : ..................................................................
C2    : ..................................................................
C3    : ..................................................................
I1    : ............C.......................T...........................
I2    : ..................................................................
P1    : ..................................................................
M1    : ...................................T.....A..C..TG........T...
U1    : ....CC..................................G........GTTT.ACAAC...
U2    : ....CC..................................G........GTTT.ACAAC...
FT1   : ....CC..................................AT.T..CA.TC.A..C..C....
FT11  : ....CC..................................AT.T..CA.TC.A..C..C....
T1    : ....CC..................................AT.T..CA.TC.A..C..C....
T11   : ....CC..................................AT.T..CA.TC.A..C..C....
        TCGCAATCCCCTATATTCATCCAACCAACCCCTTCGCCCCCGATGTCACCGCTGCGGCCG

                  *        420         *        440         *        460
B1    : ..................................................................T....
B2    : .....A.............................................................T....
C1    : .......T...........................................................
C2    : ..................................................................
C3    : ..................................................................
I1    : ...................................................................A..
I2    : ...T...............................................................
P1    : ..................................................................
M1    : .....A.....G.....T......A.......A..T...............T........
U1    : ..G..GC....C..C.....T...................T..A........T...T....
U2    : ..G..GC....C..C..T......................T...........T........
FT1   : ..G..AAT...T...........................T.................T..
FT11  : ..G..G.....C........T...........C.....AC...........A.....C..
T1    : ..G..AAT...T...........................T.................T..
T11   : ..G..G.....C........T...........C.....AC...........A.....C..
        CAACCCGCCCGACCACTCGGCTCCGCTTGGCGTGACCAGGCCCAGCGCCCCGCCGCTGCC

                  *        500
B1    : ...................... : 499
B2    : ...................... : 499
C1    : ...................... : 499
C2    : ...................... : 499
C3    : ...................... : 499
I1    : ...................... : 499
I2    : ...................... : 499
P1    : ...................... : 499
M1    : ...C.....T...G....G. : 499
U1    : .........T.......... : 496
U2    : .........T........G. : 496
FT1   : .........TT.....T.G. : 429
FT11  : .........TT.....T... : 430
T1    : .........TT.....T.G. : 494
T11   : .........TT.....T... : 500
        AGCTGGGGCCGCGCCGCTAA
```

图 1　核苷酸序列的比较

Fig. 1　Alignment of nucleotide sequences

Note: B1, Burmese prototype, Accession No. M73218[6]; B2, Burmese isolate, D10330[6,7]; P1, Pakistan isolate, M80581[8]; C1, Chinese isolate, L25547[9]; C2, Chinese isolate, M94177[13]; C3, Chinese isolate, D11093 (Uchida *et al.*, unpublished); I1, Indian isolate, X98292[10]; I2, Indian isolate, X99441 (Von Brunn *et al.*, unpublished); M1, the Mexican prototype, M74506[11] and US-1 and US-2, US isolates, AF060668 and AF060669[12].

2.5　ORF3 氨基酸同源性分析

将 T1 和 T11 的 ORF3 氨基酸序列与已报道的 11 株 HEV 的 ORF3 氨基酸序列的同源性进行了比较(见表 2), T1 和 I 型的 ORF3 的核苷酸同源性为 75.9%～77.8%; 和 II 型的同源性为 75%; 和 III 型的同源性为 79.6%～83.3%; T11 和 I 型的同源性为 76.3%～77.2%; 和 II 型的同源性为 78.1%; 和 III 型的同源性为 83.3%～86%。T1 和 T11ORF3 氨基酸的同源性为 97.7%。

维普资讯 http://www.cqvip.com

表1 基因序列的核苷酸同源性(%)比较

Table 1 Parewise comparison of nucleotide sequence (%)

	U1	U2	B1	B2	C1	C2	C3	P1	I1	I2	M1	T1
T11	85	85	81	81	81	81	81	82	82	81	81	90
T1	83	84	80	80	80	80	79	80	80	80	80	
M1	80	82	86	86	86	86	86	87	86	85		
I2	82	82	98	97	97	96	96	97	96			
I1	81	82	96	96	96	96	96	97				
P1	81	81	97	96	98	98	98					
C3	81	80	96	96	97	97						
C2	81	80	97	95	98							
C1	81	81	97	96								
B2	81	81	98									
B1	81	81										
U2	94											

Note: B1, Burmese prototype, Accession No. M73218[6]; B2, Burmese isolate, D10330[6,7]; P1, Pakistan isolate, M80581[8]; C1, Chinese isolate, L25547[9]; C2, Chinese isolate, M94177[13]; C3, Chinese isolate, D11093 (Uchida *et al*., unpublished); I1, Indian isolate, X98292[10]; I2, Indian isolate, X99441 (Von Brunn *et al*., unpublished); M1, the Mexican prototype, M74506[11] and US-1 and US-2, US isolates, AF060668 and AF060669[12].

表2 ORF3 氨基酸序列的同源性(%)比较

Table 2 Parewise comparisons (% amino acid identity) of ORF 3 sequences

	B1	B2	P1	C1	C2	C3	I1	I2	M1	U1	U2	T1
T11	77.2	76.3	77.2	77.2	76.3	76.3	77.2	77.2	78.1	86.8	83.3	87.7
T1	76.9	76.9	76.9	77.8	77.8	75.9	76.9	77.8	75.0	83.3	79.6	
U2	84.4	83.6	84.4	82.2	82.8	82.8	84.4	84.4	80.3	96.7		
U1	84.4	83.6	84.4	82.2	84.4	82.8	84.4	84.4	78.7			
M1	87.0	87.0	87.0	85.4	85.4	85.4	88.6	87.0				
I2	100	98.4	100	98.4	98.4	98.4	98.4					
I1	98.4	96.7	98.4	96.7	96.7	96.7						
C3	98.4	96.7	98.4	96.7	96.7							
C2	98.4	96.7	98.4	96.7								
C1	98.4	96.7	98.4									
P1	100	98.4										
B2	98.4											

Note: B1, Burmese prototype, Accession No. M73218[6]; B2, Burmese isolate, D10330[6,7]; P1, Pakistan isolate, M80581(8); C1, Chinese isolate, L25547[9]; C2, Chinese isolate, M94177[13]; C3, Chinese isolate, D11093 (Uchida *et al*., unpublished); I1, Indian isolate, X98292[10]; I2, Indian isolate, X99441 (Von Brunn *et al*., unpublished); M1, the Mexican prototype, M74506[11] and US-1 and US-2, US isolates, AF060668 and AF060669[12].

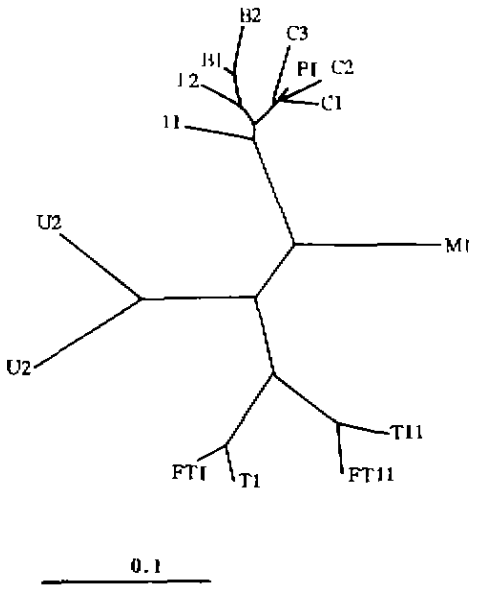

图2 基因进化树

Fig.2 Unrooted phylogenetic tree

3 讨论

国际上已对来自于不同地区的十几株戊肝病毒进行了全基因的克隆及测序，序列分析显示目前戊肝病毒可分为三个基因型，并有明显的区域性。I型来自于亚洲和非洲；II型来自于墨西哥；III型来自于美国。三型之间的核苷酸同源性为74%～76%。我们实验室克隆出IV型HEV并对其ORF2部分基因序列进行了序列分析，本文对其部分ORF1和完整的ORF3基因序列进行分析，从基因序列的同源性和基因进化树分析均显示T1和T11不同于HEV I型、II型和III型为一新的基因型，即HEV IV型。说明我国存在HEV I型和IV型两种基因型的流行。

HEV IV型与已报道的戊肝病毒相比，其基因结构也发生了改变，这是由于在该扩增区140位(全序列第5160位)核苷酸插入一个"T"(图1)，使ORF2的起始密码前移42个核苷酸，ORF3的起始

维普资讯 http://www.cqvip.com

密码后移28个核苷酸,ORF2和ORF3多肽的氨基端发生了明显变化,而氨基端主要为信号多肽,通过计算机分析显示,改变后的ORF2和ORF3仍有完整的信号多肽,说明虽然基因结构发生了改变,但仍能保证其正常的功能。而且通过对同一基因型的另一株病毒进行该片段的基因序列分析,也发现了同样的基因结构,说明这一结构为戊肝病毒IV型的特征性结构。

HEV ORF3区为GC丰富区,形成复杂的二级结构,用常规的PCR扩增易造成碱基缺失,曾报道缺失146个碱基,误认为是基因结构的改变[5],实际上是检测方法造成的。本文用常规的PCR扩增GC丰富区造成71个碱基缺失,但在常规的PCR反应液中加入GC melt溶液和二甲基亚枫可以改变复杂的二级结构,从而可扩增了GC丰富区。

参考文献

[1] Harrison TJ Hepatitis E virus: an update[J]. Liver, 1999, 19: 171-176.

[2] Wang Youchun, Ling R, Erker J C, *et al*. A divergent genotype of hepatitis E virus in Chinese patients with acute hepatitis[J]. J Gen Viro 1999, 80, 169-177.

[3] 杨军,张华远,王佑春,等.三株戊型肝炎病毒变异株的核苷酸比较[J].中国病毒学,(刊印中).

[4] 王佑春,Morris A, Dashiko G,等.庚型肝炎病毒(HGV/GBV-C)NS3区基因序列的多变性[J].中华微生物学和免疫学杂志,1997,17(2):96-98.

[5] Ray R, Jameel S, Manivel V, *et al*. Indian hepatitis E virus shows a major deletion in the small open reading frame[J]. Virology 1992, 189:359-362.

[6] Tam A W, Smith M M, Guerra M E, *et al*. Hepatitis E virus (HEV): molecular cloning and sequencing of the full-length viral genome[J]. Virology, 1991, 185: 120-131.

[7] Aye T T, Uchida T, Ma, *et al*. Sequence and gene structure of the hepatitis E virus isolated from Myanmar [J]. Virus Genes, 1993, 7: 95-110.

[8] Tsarev S A, Emerson S U, Reyes G R, *et al* Characterization of a prototype strain of hepatitis E virus [J]. Proc Nat Acad Sci U S A, 1991, 89: 559-563.

[9] Bi S L, Purdy M A, McCaustland, *et al*. The sequence of hepatitis E virus isolated directly from a single source during an outbreak in China [J]. Virus Research. 1993, 28, 233-247

[10] Donati M C, Fagan E A, Harrison T J. Sequence analysis of full length HEV clones derived directly from human liver in fulminant hepatitis E[M]. In: Viral Hepatitis and Liver Disease (Edited by Rizzetto, M. *et al*.), 1997: 313-316.

[11] Huang C C, Nguyen D, Fernandez J, *et al*. Molecular cloning and sequencing of the Mexico isolate of hepatitis E virus (HEV) [J]. Virology 1992, 191: 550-558.

[12] Schlauder G G, Dawson G J, Erker J C, *et al*. The sequence and phylogenetic analysis of a novel hepatitis E virus isolated from a patient with acute hepatitis reported in the United States [J]. J Gen Virol, 1998, 79: 447-456.

[13] Yin S R, Purcell R H, Emerson S U. A new Chinese isolate of hepatitis E virus: comparison with strains recovered from different geographical regions [J]. Virus Genes, 1994; 9, 23-32