

## HGTV 广东株和香港株 NS5 区部分基因的克隆及序列分析\*

马会慧<sup>1</sup> 杨绍基<sup>1</sup> 李刚<sup>1</sup> 陈雪娟<sup>1</sup> 廖家杰<sup>2</sup> 姚集鲁<sup>1</sup><sup>1</sup> (中山医科大学传染病学教研室, 广州 510630)<sup>2</sup> (香港大学玛丽医院胃肠肝病科, 香港)

12373-21

**摘要** 采用 HGV NS5 特异的 2 对引物, 对两个香港株和一个广东株 HGV-RNA 进行逆转录套式 PCR 扩增, PCR 产物克隆入 pUC19, 重组质粒转化 DH5 $\alpha$  和 JM109 菌株。PCR 和酶切法鉴定阳性克隆, 双脱氧链末端终止法测定核苷酸序列并进行同源性分析。结果发现核苷酸变异呈散在分布, 三株间核苷酸和氨基酸序列同源性分别为 93.3%~94% 及 97%~99.2%, 与已报道的中国株 (CN) 相比, 则同源性分别为 90%~91.2% 和 94%~96.3%, 与美国株 (PNF2161 及 R10291) 相比, 为 87.1%~89.5% 和 95.2%~97%, 而与西非株 (GBV-C) 相比, 则达 91.4%~93.8% 和 97%~97.9%。提示 HGV NS5 区核苷酸和氨基酸序列相对保守, 不同 HGV 株存在一定的地区差异。

**关键词** 庚型肝炎病毒, 聚合酶链反应, DNA 序列测定, 分子克隆

广东株 香港株

PCR

庚型肝炎病毒 (HGV) 是两年前发现的一种新型肝炎病毒, 其基因组包括 5'-非编码区 (5'-NTR), 3'-非编码区 (3'-NTR), 结构蛋白 C、E1、E2 区, 有的株 C 区缺如, 非结构蛋白 NS1、NS2、NS3、NS4A、NS4B、NS5A/5B 区<sup>[1]</sup>。国内外对 5'-NTR、E2 区、NS3 区等的研究报道较多, 但对 NS5 区的报道甚少。为了解中国不同地区之间 HGV NS5 区基因以及与国外分离株的同源性状况, 本研究对 HGV NS5 区部分基因进行了 PCR 克隆及序列分析, 现报道如下。

## 1 材料与方法

## 1.1 血清标本

两份血清来自香港献血员 (A132、A711), 一份血清来自广东的一名慢性乙型肝炎病人 (3027)。

## 1.2 主要试剂

RNasin、AMV 逆转录酶、Taq DNA 聚合酶 (购自丰华公司)、pUC19、IPTG、X-gal (天象人生物工程公司)、低熔点琼脂糖、T4DNA 连接酶、BamHI、SmaI、EcoRI、Klenow 酶、琼脂酶 (宝灵曼公司)。

引物由北京赛百胜生物工程公司合成, 序列及位置如下:

33 5'-GTTGAATTCGCGATGGAGCGCTACAC3' (6672~6697)

34 5'-CTGGGATCCGTATCATGTATGGTTCT3' (7267~7292)

35 5'-TCGATTGCTGTAGCTGAGCC3' (6573~6592)

36 5'-GGTAAGTTCATTGCCACCA3' (7327~7346)

收稿日期: 1998-03-05, 修回日期: 1998-05-15

\* 本课题由国家自然科学基金 (编号 39 600 130) 和广东省医学科学研究基金部分资助

33、34 为一对内引物, 35、56 为一对外引物。单划线部分为 BamHI 位点, 双粗划线部分为 EcoRI 酶切位点。外引物扩增片段为 774 bp, 内引物扩增片段为 618 bp。

### 1.3 PCR 产物的获取

采用热变性酶裂解法制备 RNA 模板<sup>[2]</sup>。取待检标本 50  $\mu$ L, 加入裂解液 5  $\mu$ L (主要含蛋白酶 K), 55  $^{\circ}$ C 32 min, 98  $^{\circ}$ C 15 min, 12 000 g/min 离心 5 min。取裂解上清 5  $\mu$ L, 进行逆转录, 反应体积 10  $\mu$ L, 含引物 36、75 ng, AMV 11 u, 2.5 mmol/L dNTP 1.5  $\mu$ L, Rnasin 16 u, 42  $^{\circ}$ C 60 min。第一次 PCR 反应体积为 20  $\mu$ L, 含逆转录液 10  $\mu$ L, 引物 35 为 50 ng, Taq DNA 聚合酶 2 u, 95  $^{\circ}$ C 40 s, 55  $^{\circ}$ C 40 s, 72  $^{\circ}$ C 1 min 20 s, 35 个循环。第二次 PCR 反应体积为 20  $\mu$ L, 内含第一次 PCR 反应液 5  $\mu$ L, 引物 33 为 75 ng, 引物 34 为 75 ng, Taq DNA 聚合酶 2 u, 95  $^{\circ}$ C 30 s, 60  $^{\circ}$ C 30 s, 72  $^{\circ}$ C 40 s, 30 个循环。取 10  $\mu$ L 在含溴化乙锭的 2% 的琼脂糖凝胶中电泳, 与标准分子量作对照, 紫外灯下观察结果。

### 1.4 PCR 产物的分子克隆

A132、A711 的 PCR 产物用 Klenow 酶补平, 1% 低熔点琼脂糖电泳, 切下目的条带处的凝胶, 琼脂酶消化, 乙醇沉淀, 回收 DNA。用 BamHI 消化, 与用 SmaI 和 BamHI 消化的 pUC19 连接。连接反应总体积 20  $\mu$ L, 内含回收 DNA 10  $\mu$ L, pUC19 6  $\mu$ L, 10 $\times$  连接 Buffer 2  $\mu$ L, T4DNA 连接酶 2 u, 14  $^{\circ}$ C 过夜。加入 80  $\mu$ L DH<sub>5</sub> $\alpha$  受体菌中。而 3027 株则用 BamHI 和 EcoRI 消化, 克隆入 pUC19 的 BamHI 和 EcoRI 位点, 转化 JM<sub>109</sub> 受体菌。受体菌制备及转化条件按文献[3]的方法。碱变性法提取质粒, PCR 及单双酶切鉴定阳性克隆。

### 1.5 DNA 序列测定

重组的阳性克隆, pUC19 模板采用双脱氧链末端终止法, 以 DNA 全自动测序仪 (ABI373) 测定核苷酸序列, 同一片段经正反两个方向重复测定。

## 2 结果

### 2.1 阳性克隆的鉴定

A132、A711 的 HGV NS5 区的 618 bp 的片段插入 pUC19 的 SmaI 和 BamHI 之间, 3027 的 NS5 区的 425 bp 的片段插入在 pUC19 的 EcoRI 和 BamHI 之间。重组质粒 pUGN<sub>132</sub>、pUGN<sub>711</sub> 以及 pUGN<sub>3027</sub> 经 BamHI 单酶切后与空载体 pUC19 比较电泳时泳动较慢。pUGN<sub>132</sub>、pUGN<sub>711</sub> 经 BamHI 和 SacI 酶切, 则切出一条带, 与插入片段相符, 用特异引物 33、34 作 PCR 扩增, 获得 618 bp 产物。pUGN<sub>3027</sub> 经 EcoRI 和 BamHI 酶切, 则切出一条 425 bp 条带, 与插入片段相符。用上述克隆进行序列测定。

### 2.2 HGV 的 A132、A711、3027 NS5 区部分片段的核苷酸序列及与其他分离株的比较 见图 1、图 2 及表 1。

## 3 讨论

HGV 与 HCV 相比, 其基因组变异小, 不存在高变区, 同一基因组不同区段变异程度不同<sup>[6]</sup>, 也存在地区差异<sup>[7]</sup>。5'-非编码区及 3'-非编码区高度保守, 结构蛋白区变异较大, 非结构蛋白区存在保守基序, 其中的 NS5 区报道很少。本研究对两个香港地方株和一个广东株的 NS5 区部分基因进行了 PCR 克隆及序列测定分析。三株序列间相比, 该段核苷酸的变异小, 同源性大于 93%, 而氨基酸同源性高达 97% 以上; 三株序列各与 CN 株相比, 核苷酸和推导的氨基酸序列的同源性分别为 90%~91.2% 和 94%~96.3%; 而与美国株相比同源性略低, 分

A132	gttgaattcgggatggagcgtacacTCTTCCGCGCCAACTGCCGATGAGGAACGTGGCG	60
A711	C..C....A...GT.....T.....C	
PNF	CT.G..T.AT.....A..C.C....T....A	
R10291	C..G....AT.....C.C....T..A...	
GBV-C	C.....A...GT.....C...T.....	
CN	C.....A.....	
A132	CCCTCTGAGGTTTCATCTGAGGTCAGCATCGAGATCGGGACGGAGACTGAGGATTCAGAA	120
A711	..T.....G..A.....T.....A..C.....	
PNF	.....C....GTC..T..C..T.....A..C.....	
R10291	.....C....GTC..A..C..T.....A..C.....	
GBV-C	.....A..C.....	
CN	.....G.....T.....A..C.....	
A132	CTGACTGAGGCAGATCTGCCACCCGCAGCTGCCGCCCTTCAGGCGATCGAGAATGCTGCC	180
A711	.....T.....T.....T..C..A....A.....	
PNF	.....C.....G..G..G....T..T..C..A.....	
R10291	.....C..C....G..G....A....C....T.....	
GBV-C	.....C..T.....A..G....T....C..A....A.....	
CN	.....A.....	
A132	AGAATCTTGAGCCGCACATTGATGTCATCATGGAGGACTGCAGTACACCCTCTCTCTGT	240
A711	.....C..A.....	
PNF	..G.....A.....T...	
R10291	..G.....T..T.....T.....T...	
GBV-C	.....C..A.....C.....Y.....T.....	
CN	.....A.....T...	
3027	.....T.....A.....	
A132	GGGAGTAGCCGAGAGATGCCTGTGTGGGGAGAAGACGTACCCCGCACTCCATCGCCTGCA	300
A711	..T.....A.....	
PNF	..T.....A.....A.C....T.....A...	
R10291	..T.....A.C.....A...	
GBV-C	..T.....A.....	
CN	..T.....A.....A.C.....A...	
3027	..A.....A.....C.....A.....	
A132	CTTATCTCGGTTACGGAGAGCAGCTCAGATGAGAAGACCCCGTCGGTGTCTCTTCGCAG	360
A711	.....	
PNF	.....T.....C.....	
R10291	.....C.....C.....	
GBV-C	.....T.....A...C.....	
CN	.....C.....T.....C.....	
3027	.....C.....	

A132	GAGGATACCCCGTCCCTCGGACTCATTGAAAGTCATCCAAGAGTCTGATACTGCTGAGAGT	420
A711	..... A. G. .... G. ....	
PNF	..... T. .... C. . G. .... C. . G. . A. . C. . AG. G	
R10291	..... T. .... C. .... G. . A. .... AG. A	
GBV-C	..... C. .... A. .... ATCA	
CN	..... C. .... T. .... G. .... C. . G. . A. . C. . A. CG	
3027	.....	
A132	GAGGAAAGCGTCTTCAACGTGGCTCTTCCGTAATAAAAGCCTTATTTCCACAAAGCGAT	480
A711	..... T. .... C. .... G. ....	
PNF	..... T. .... T. .... G. .... C	
R10291	..... T. .... G. .... G. .... G. . T. .	
GBV-C	..... G. ....	
CN	..... G. .... C. .... G. .... C	
3027	..... C. . G. . G. ....	
A132	GCTAAGCGTAAGCTCACAGTCAAGATGTCATGCTGTGTGGAGAAGAGCGTCAAGCGCTTC	540
A711	.. C. .... T. . T. .... A. ....	
PNF	.. G. . CA. G. .... T. . C. .... G. .... C. . T. . A. .... T	
R10291	.. C. . TA. A. .... T. . C. .... G. .... AAT. .... C. . T. ....	
GBV-C	.. C. . A. . A. .... A. . G. . T. .... T. .... T. .... A. . A. ....	
CN	.. G. . TA. G. .... T. . C. .... G. .... C. . T. . A. .... T	
3027	.. C. . A. .... A. . G. . T. . G. .... A. ....	
A132	TTTTCCCTTAGGGTTGACTGTAGCTGATGTGGCTAGCCTGTGTGAGATGGAAATCCagaaccatacatgatacggatcc	618
A711	..... A. . G. .... TC. A. ....	
PNF	.. C. . A. . G. .... G. . G. .... T. .... C. ....	
R10291	..... T. . G. . C. .... G. . G. .... C. . TC. .... G. ....	
GBV-C	..... T. .... C. . G. .... C. .... C. .... G. ....	
CN	.. C. . A. . G. .... G. . G. .... T. .... C. .... G. ....	
3027	..... C. .... C. . G. . C. . C. .... C. .... A. ....	

图 1 HGV A132(香港株)NS5 cDNA 序列与 A711(香港株)、3027(广东株)及已报道的 CN<sup>[4]</sup>、PNF2161<sup>[11]</sup>、R10291<sup>[12]</sup>和 GBV-C<sup>[3]</sup>分离株相应序列的比较。小写为引物序列。圆点表示与 HGV A132 株序列相同。

Fig. 1 Comparison of NS5 cDNA sequence of HGV A132 (Hong Kong strain) with the corresponding sequences from 3027 (Guangdong strain), A711 (Hong Kong) and the reported isolates (CN, PNF2161, R10291, GBV-C). Small letters stand for the sequence of primers. Dots indicate identity with sequence of HGV A132

别为 87.1%~89.5% 和 95.2%~97%;有趣的是,与 GBV-C 相比,核苷酸的同源性为 91.4%~93.8%,氨基酸的同源性则达 97%~97.9%。结果提示 HGV NS5 区的该段基因序列存在一定的地区差异,支持 HGV 基因变异可能存在地域因素的观点。但总体而言,HGV NS5 区的该段基因序列相对保守,7 株间相比,核苷酸和氨基酸的同源性分别达 87.1%~94% 和 94%~99.2%。核苷酸变异主要呈散在分布,但以 6698~6837 bp、7080~7092 bp 和 7157~7251 bp(以 PNF2161 计位)较为集中。

A132	LPRQLRMRNV APSEVSSEVS IEIGTETEDS ELTEADLPPA AAALQAIENA ARILEPHIDV	60
A711	..H.....	
PNF	..H...L... ..D.....	
R10291	..H...L... ..D.....	
GBV-C	..H.....	
CN	..H.....	
3027	.....	
A132	IMEDCSTPSL CGSSREMPVW GEDVPRTPSP ALISVTESSS DEKTPSVSSS QEDTPSSDSF	120
A711	.....I.....G..	
PNF	.....I.....	
R10291	.....I.....	
GBV-C	.....I.....L..T..	
CN	.....I.....L.....L	
3027	.....	
A132	EVIQESDTAE SEESVFNVAL SVLKALFPQS DATRKLTVKM SCCVEKSVTR FFSGLGLTVAD	180
A711	....G.....	
PNF	....E... G.....	
R10291	....E... G.....E.....R. N.....	
GBV-C	.....	
CN	....E... T.....E.F.....	
3027	.....R.....	
A132	VASLCEMEI	189
A711	.....	
PNF	.....	
R10291	.....	
GBV-C	.....	
CN	.....	
3027	.....	

图 2 HGVA132 NS5 区推导的氨基酸序列与 A711、3027 株及 CN、PNF2161、R10291、GBV-C 相应序列的比较  
 Fig. 2 Comparison of deduced amino acid sequence of HGVA132 NS5 with the corresponding sequences from 3027, A711 strains and CN, PNF2161, R10291, GBV-C isolates

表 1 HGVA132、A711、3027 株与其他分离株之间的 NS5 区部分基因同源性比较

Table 1 Homology analysis of HGVA NS5 sequence from different strains

nt/aa*	A132	A711	3027	PNF	R10291	GBV-C
A711	93.3/97.9	-	-	-	-	-
3027	94/99.2	93.8/97	-	-	-	-
PNF	87.5/96.8	87.5/96.8	88.3/97	-	-	-
R10291	88/95.8	87.1/95.2	89.5/96.2	92.3/98.4	-	-
GBV-C	91.4/97.9	93.8/97.9	93.3/97	87.9/97.9	88.9/95.2	-
CN	91.2/95.8	90/96.3	90/94	93.3/96.3	90.9/95.8	90.7/96.8

\* nt = nucleotide, aa = amino acid

本研究中 3027 株所测序列仅有 402 bp, 比设想的片段小 167 bp, 是因为在该段核苷酸的 5' 端第 182 bp 处存在一个 EcoRI 酶切位点, 当以 EcoRI 和 BamHI 酶切时, 则仅有 425 bp 接入 pUC19 的 EcoRI 和 BamHI 之间, 这是设计引物应注意的问题。已报道的 4 株 HGV 中, CN 和 GBV-C 株在该处存在一个 EcoRI 酶切位点, 而 PNF2 161 和 R10 291 则无, 这也与我们所测序列与 CN 和 GBV-C 株的同源性高相符。

### 参 考 文 献

- 1 Linnen J, Wages J, Zhang-Keck ZY *et al*. Molecular cloning and disease association of hepatitis G virus: a transfusion-transmissible agent. *Science*, 1996, 271:505-508
- 2 高志良, 姚集鲁 热变性 HCV RNA 模板直接法扩增 中山医科大学学报, 1994, 15(3):68
- 3 Sambrook J, Fritsch EF, Maniatis T *et al*. *Molecular cloning. a laboratory Manual*. Cold Spring Harbor Laboratory Press, 1989
- 4 周育森, 陈薇, 赵秋敏等. 中国人庚型肝炎病毒全基因 cDNA 克隆及序列测定. 军事科学院院刊, 1996, 20(4):249-254
- 5 Leary TP, Muerhoff AS, JN simons JN *et al*. Sequence and genomic organization of GBV-C: a novel member of the *Flaviridae* associated with human non A-E hepatitis. *Journal of Medical Virology*, 1995, 48:60-67
- 6 Erker JC, Simons JN, Muerhoff AS *et al*. Molecular cloning and characterization of a GB virus C isolate from a patient with non-A-E hepatitis. *Journal of General Virology*, 1996, 77:2713-2720
- 7 Kao JH, Chen PJ, Hsiung SC *et al*. Phylogenetic analysis of GB virus C: comparison of isolates from Africa, North America, and Taiwan. *Journal of Infectious diseases*, 1996, 174:410-413

## Molecular Cloning and Sequencing the NS5 Genomic Region of HGV Guangdong and Hong Kong Strains

Ma Huihui Yang Shaoji Li Gang *et al*

(*Department of Infectious Diseases, Sun Yat-Sen University of Medical Sciences, Guangzhou 510630*)

**Abstract** HGV NS5 cDNA of two Hong Kong strains and one Guangdong strain were amplified by reverse transcription polymerase chain reaction (RT-PCR). The products were inserted into pUC19 vectors respectively. After transfecting DH5 $\alpha$  and JM109 strains, the recombinants were screened and identified by PCR and digested with restriction endonucleases. The nucleotides were sequenced by the dideoxy chain termination method. The homologies of HGV NS5 region nucleotide and deduced amino acid sequences were 93.3% - 94% and 97% - 99.2% respectively between each Hong Kong strain and the Guangdong strain, and were 90% - 91.2% and 94% - 96.3%, 87.1% - 89.5% and 95.2% - 97%, 91.4% - 93.8% and 97% - 97.9% respectively compared with the reported China isolate (CN), American isolates (PNF2161 and R10291), and West African isolated (GBV-C). The variant sites of the nucleotides appear sporadically. The nucleotide and deduced amino acid sequences of HGV NS5 genomic region are relatively conserved, and it also suggests that the nucleotide variation of HGV may be influenced by geographical factors.

**Key words** HGV, polymerase chain reaction, DNA sequence, Molecular cloning