

# 乙型肝炎病毒基因组前-前-S基因区的分子流行病学研究

杨倩,董菁,成军

杨倩,董菁,成军,中国人民解放军第302医院传染病研究所基因治疗研究中心、全军病毒性肝炎防治研究重点实验室 北京市 100039  
杨倩,女,1970-11-06,广西桂平人,汉族,1999年毕业于西安医科大学,获硕士学位,目前在西安交通大学第一医院攻读博士研究生,主要从事传染病临床与基础研究工作。  
国家自然科学基金攻关项目, No. C03011402, No. C30070689  
军队“九、五”科技攻关项目, No. 98D063  
军队回国留学人员启动基金项目, No. 98H038  
军队“十、五”科技攻关青年基金项目, No. 01Q138  
军队“十、五”科技攻关面上项目, No. 01MB135  
项目负责人:成军,100039,北京市西四环中路100号,中国人民解放军第302医院传染病研究所基因治疗研究中心、全军病毒性肝炎防治研究重点实验室。cj@genetherapy.com.cn  
电话:010-66933392 传真:010-63801283  
收稿日期:2003-11-13 接受日期:2003-12-16

## Molecular epidemiological features of pre-pre-S region in hepatitis B virus genome

Qian Yang, Jing Dong, Jun Cheng

Qian Yang, Jing Dong, Jun Cheng, Gene Therapy Research Center, Institute of Infectious Diseases, The 302 Hospital of PLA, Beijing 100039, China  
Supported by Grants from National Natural Science Foundation No. C03011402, No. C30070689, and Returned Scholarship of General Logistics Department of PLA.  
Correspondence to: Dr. Jun Cheng, MD, PhD, Director, Gene Therapy Research Center, Institute of Infectious Diseases, The 302 Hospital of PLA, 100 Xisihuanzhong Road, Beijing 100039, China. cj@genetherapy.com.cn  
Received: 2003-11-13 Accepted: 2003-12-16

### Abstract

**AIM:** To try to identify pre-pre-S region in hepatitis B virus (HBV) genome and to study the relationship between the genotype and the existence of pre-pre-S region.

**METHODS:** Genomic DNA of HBV was extracted from the sera of patients with HBV infection. Multiple-primers polymerase chain reaction (PCR) method was used to identify the genotype of HBV strains. Then the pre-pre-S region was amplified by PCR method, and TA cloned into pGEM Teasy vector. After sequencing of the target region, vector 8.0 software was used to analyse the sequences.

**RESULTS:** Fifteen male patients were collected in this study. There were 1 genotype B, 5 genotype B and C mixture, and 9 genotype C. Amplified PCR products of the pre-pre-S region from these 15 samples were ligated into the pGEM Teasy vectors. After sequencing, 31 clones were found being encoding for pre-pre-S peptide.

**CONCLUSION:** Existence of coding gene of the pre-pre-S region is popular in B or C genotype of HBV.

Yang Q, Dong J, Cheng J. Molecular epidemiological features of pre-pre-S region in hepatitis B virus genome. *Shijie Huaren Xiaohua Zazhi* 2004;12(4):785-789

### 摘要

**目的:** 探讨乙型肝炎病毒(HBV)基因组是否存在前-前-S编码基因,并探讨前-前-S基因与HBV基因型分布之间的关系。

**方法:** 自慢性乙型肝炎患者血清中提取病毒基因组DNA,首先应用多引物-多聚酶链反应(PCR)进行HBV基因型分型,之后将前-前-S区扩增,TA克隆到pGEM Teasy载体后进行单克隆测序,利用Vector 8.0软件进行序列分析。

**结果:** 选择15例患者,经过HBV基因型分型实验确定基因型为B型者1例,B/C混合型者5例,C型9例。自15例患者血清中提取的HBV基因组中扩增出前-前-S基因片段,TA克隆后选择31个克隆进行测序,测序结果证明这31个克隆均编码前-前-S基因。

**结论:** 前-前-S区编码在B、C基因型中均为普遍存在的现象。

杨倩,董菁,成军.乙型肝炎病毒基因组前-前-S基因区的分子流行病学研究. *世界华人消化杂志* 2004;12(4):785-789  
<http://www.wjgnet.com/1009-3079/12/785.asp>

### 0 引言

乙型肝炎病毒(HBV)基因组全长在3200个核苷酸(nt)左右,为部分双链DNA病毒。Gelibert et al<sup>[1-2]</sup>于1979年发表了HBV基因组的第一个全长核苷酸序列,我国学者<sup>[3-4]</sup>于1984年报道了中国大陆HBV全序列。Gelibert et al最初的研究将HBV基因组划分出4个开放读码框架(ORF)<sup>[1]</sup>,分别命名为S、C、P、X区,4个ORF中表达的氨基酸长度不同,其生物学功能也不相同<sup>[5]</sup>,其中全S区又因不同的起始密码子(ATG)而又人为的分为前-S1、前-S2和S三个区,前-S1、前-S2和表面抗原主蛋白是按照同一开放读码顺序(in frame)进行翻译的。我们<sup>[6-9]</sup>研究了乙型肝炎患者血清中存在的HBV病毒基因组,寻找到新的ORF存在的证据<sup>[7-9]</sup>,并初步重新确定了全S基因(前-前-S、前-S1、前-S2和主蛋白基因)和全X(前-X、X基因)基因的启动子<sup>[10-11]</sup>。本文以分子流行病学方法证实了前-前-S区的存在状态。

### 1 材料和方法

**1.1 血清采集和DNA提取** 血清来源:15例患者,诊断为病毒性肝炎,乙型,慢性,符合2000年《病毒性肝炎防治方案》(试行)<sup>[12]</sup>,临床初步检测HBV DNA阳性。DNA提取:自静脉中采集5 ml全血,分离血清,200 μl

血清加入蛋白酶K及其缓冲液, 65 °C消化过夜, 饱和酚: 氯仿(1:1)抽提总DNA, 无水乙醇沉淀, 加入无DNA酶的RNA酶消化后, -20 °C保存备用。

**1.2 HBV基因型分析** 参考Naito et al<sup>[13]</sup>设立的6种主要基因型分型方法, 建立HBV基因型的检测方法。简言之, 应用巢式-多引物-多聚酶链反应, 外引物: 上游引物序列: 5' - AGC ATG GGA GGT TGG TCT TC-3', 下游引物序列: 5' - AAG GCA TCA AGG CAG GAT AGC-3', 目的片段长度约1 429 bp, 自前-S1区起始处至S区终止子下游226 bp。内引物: 上游引物序列: 5' - GGC TCA AGT TCC GGA ACA GT-3', A基因型特异性下游引物: 5' - CTC GCG GAG ATT GAC GAG ATG T-3', B基因型特异性下游引物: 5' -CAG GTT GGT GAG TGA CTG GAG A-3', C基因型特异性下游引物: 5' - GGT CCT AGG AAT CCT GAT GTT G-3'。外引物扩增PCR程序参数如下: 94 °C 1 min 预变性, 94 °C 1 min 30 s 变性, 59 °C 1 min 30 s 退火, 72 °C 2 min 钟延长, 共30个循环, 72 °C延长10 min。内引物扩增PCR程序参数如下: 94 °C 1 min 预变性, 94 °C 30 s 变性, 58 °C 30 s 退火, 72 °C 30 s 延长, 共35个循环。NT-PCR产物经过3%琼脂糖凝胶电泳后判断HBV基因型。

**1.3 前-前-S区的克隆及序列分析** 采用本研究组以往报告<sup>[14-16]</sup>的PCR-TA克隆-测序方法。多聚酶链反应(PCR)扩增前-前-S区: 前-前-S上游引物为P1: 5' - CTA GAT TCT ATC CTA CCC ACA C', P2: 5' - TTG GGA ACA AAA GAT TCG TCC-3', P1下游第70-72 nt处为前-前-S区起始密码子ATG, P2上游30-32 nt处为前-S1区起始密码子ATG。PCR参数如下: 94 °C 1 min 预变性, 94 °C 60 s 变性, 60 °C 60 s 退火, 72 °C 60 s 延长, 共35个循环, 72 °C再延长10 min。

**克隆目的片段:** 将PCR产物在1%琼脂糖凝胶中电泳, 切取目的片段, 玻璃奶法回收PCR产物, 与Promega公司所产的pGEM Teasy载体连接过夜。将连接好的重组质粒转入细菌JM109, 氨苄青霉素(Amp)和X-gal 蓝白斑法筛选阳性菌落, 进一步提取质粒鉴定。DNA测序: 选择经鉴定pGEM Teasy内插入约220 bp产物的菌落送检测序, 由上海博亚公司完成。测序引物为pGEM Teasy载体自有的T7、SP6测序引物。

**1.4 序列分析** 应用DNASIS 2.5和Vector 8.0版软件对进行了测序分析的HBV前-前-S基因序列进行碱基序列和氨基酸序列一致性分析。该软件比较后提供两个数据: 阳性率和一致率, 阳性率是选定区域推定序列的核苷酸(氨基酸)序列数目与区域总长度之比, 推定序列是软件自动比较所有序列, 参考每个核苷酸(氨基酸)位点上每个克隆的编码方式, 由大多数克隆(多于70%)编码的核苷酸(氨基酸)组成的一致性序列, 阳性率用于展示区域内部的插入突变或/和缺失突变, 以及较少见的单一位点多种核苷酸(氨基酸)编码方式; 一致率是选定

区域全部克隆均为一致的核苷酸(氨基酸)数目与最长的单一克隆核苷酸(氨基酸)序列数目之比, 表示该段区域核苷酸(氨基酸)序列一致性, 用于展示区域内部的替换突变或/和缺失突变。各区段之间的划分是在计算机软件比较的基础上, 人为划定的, 最短的分区至少大于200 nt, 不选择阳性率小于97%的区域, 因为这提示该段内有较长的缺失突变。

## 2 结果

**2.1 经过分型研究, 获得的内引物PCR产物电泳结果见图1。**

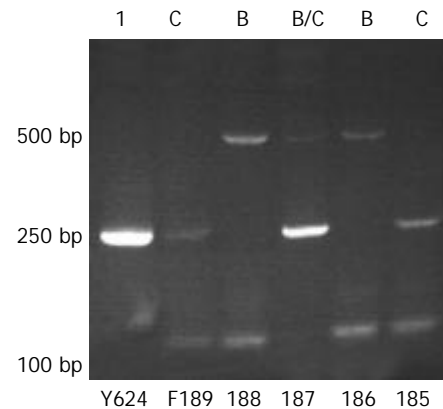


图1 HBV DNA分型的内引物PCR产物电泳结果。

应用该基因型分型方法, 电泳时展示的HBV A基因型PCR产物长度为68 bp, B型为281 bp, C型为122 bp。本研究的15例患者中, 基因型为B型者1例, B/C混合型者5例, C型9例。

**2.2 前-前-S基因核酸序列分析** 前-前-S基因全长135 bp, 31株核苷酸序列比较见图2。本研究扩增的靶区域全长218 nt, 包括前-前-S基因上游46 nt和下游前-S1区37 nt, 靶区域除去引物区比较提示阳性率为100.0%, 一致率为73.5%; 前-前-S基因编码框架自第70-204 nt, 阳性率为100.0%, 一致率为89.6%, 相对较为保守; 前-前-S基因的一致率高于靶区域的一致率, 其原因在于有3株克隆在前-S1区域出现缺失突变, 导致一致率的下降。

在序列比较中, 我们发现来自同一患者体内的克隆之间的一致性大于95%, 符合准种表现。图2中显示第20 nt存在替换突变的现象, 来自其他患者编码密码子为TCA, 而来自F244患者的3株克隆密码子为TTA (C→T替换突变, 箭头前核苷酸代表其他14个患者的编码核苷酸, 箭头后核苷酸代表某患者的特异性变异, 下同), 表现出替换突变的特异性; 此外, Y948患者2个克隆在114位点的C→T替换突变、155位点的C→A替换突变, F332患者2个克隆在128位的C→T替换突变, Y1001患者2个克隆在167位点的T→C突变, 均展示出个体化突变的现象。

		70		119
F238-2 B	(69)	ATGCAGTTAATCATTACTTCAA	AACTAGGCATTATTTACATA	CTGTGG
F238-3 B	(69)	ATGCAGTTAATCATTACTTCAA	AACTAGGCATTATTTACATA	CTGTGG
F244-1 B/C	(70)	ATGCAGTTAATCATTACTTTAAA	AACTAGGCATTATTTACATA	CTGTGG
F244-2 B/C	(69)	ATGCAGTTAATCATTACTTTAAA	AACTAGGCATTATTTACATA	CTGTGG
F244-3 B/C	(70)	ATGCAGTTAACCATTACTTTAAA	AACTAGGCATTATTTACATA	CTGTGG
F246-2 C	(70)	ATGCAGTTAATCATTACTTCAA	AACTAGGCATTATTTACATA	CTGTGG
F246-3 C	(70)	ATGCAGTTAATCATTACTTCAA	AACTAGGCATTATTTACATA	CTGTGG
F246-4 C	(70)	ATGCAGTTAATCATTACTTCAA	AACTAGGCATTATTTACATA	CTGTGG
F248-1 C	(69)	ATGCAGTTAATCATTACTTCAA	AACTAGGCATTATTTACACA	CTGTGG
F248-2 C	(70)	ATGCAGTTAATCATTACTTCAA	AACTAGGCATTATTTACATA	CTGTGG
F248-4 C	(70)	ATGCAGTTAATCATTACTTCAA	AACTAGGCATTATTTACATA	CTGTGG
F222-5 C	(70)	ATGCAGTTAATCATTACTTCAA	AACTAGGCATTATTTACATA	CTGTGG
Y948-1 C	(70)	ATGCAGTTAATCATTACTTCAA	AACTAGGCATTATTTACATA	CTTTGTGG
Y948-2 C	(70)	ATGCAGTTAATCATTACTTCAA	AACTAGGCATTATTTACATA	CTTTGTGG
Y949-1 C	(70)	ATGCAGTTAATCATTACTTCAA	AACTAGGCATTATTTACATA	CTGTGG
Y949-2 C	(69)	ATGCAGTTAATCATTACTTCAA	AACTAGGCATTATTTACATA	CTGTGG
Y951-1 C	(69)	ATGCAGTTAATCATTACTTCAA	AACTAGGCATTATTTACATA	CTGTGG
Y951-2 C	(69)	ATGCAGTTAATCATTACTTCAA	AACTAGGCATTATTTACATA	CTGTGG
Y967-1 C	(69)	ATGCAGTTAATCATTACTTCAA	AACTAGGCATTATTTACATA	CTGTGG
Y967-2 C	(70)	ATGCAGTTAATCATTACTTCAA	AACTAGGCATTATTTACATA	CTGTGG
Y991-2 C	(70)	ATGCAGTTAATCATTACTTCAA	AACTAGGCATTATTTACATA	CTGTGG
Y991-3 C	(70)	ATGCAGTTAATCATTACTTCAA	AACTAGGCATTATTTACATA	CTGTGG
Y992-1 C	(70)	ATGCAGTTAATCATTACTTCAA	AACTAGGCATTATTTACATA	CTGTGG
Y992-2 C	(70)	ATGCAGTTAATCATTACTTCAA	AACTAGGCATTATTTACATA	CTGTGG
F332-1 B/C	(70)	ATGCAGTTAATCATTACTTCAA	AACTAGGCATTATTTACATA	CTGTGG
F332-2 B/C	(70)	ATGCAGTTAATCATTACTTCAA	AACTAGGCATTATTTACATA	CTGTGG
Y999-1 B/C	(70)	ATGCAGTTAATCATTACTTCAA	AACTAGGCATTATTTACATA	CTGTGG
Y1001-1 B/C	(70)	ATGCAGTTAATCATTACTTCAA	AACTAGGCATTATTTACATA	CTGTGG
Y1001-3 B/C	(70)	ATGCAGTTAATCATTACTTCAA	AACTAGGCATTATTTACATA	CTGTGG
Y994-1 B/C	(70)	ATGCAGTTAATCATTACTTCAA	AACTAGGCATTATTTACATA	CTGTGG
Y994-2 B/C	(70)	ATGCAGTTAATCATTACTTCAA	AACTAGGCATTATTTACATA	CTGTGG
Consensus	(70)	ATGCAGTTAATCATTACTTCAA	AACTAGGCATTATTTACATA	CTGTGG
		120		169
F238-2 B	(119)	AAGGCTGGCATTCTATATAAG	AAGAGAACTACACGCAGCGC	CTCATTTTG
F238-3 B	(119)	AAGGCTGGCATTCTATATAAG	AAGAGAACTACACGCAGCGC	CTCATTTTG
F244-1 B/C	(120)	AAGGCTGGCATTCTATATAAG	AAGAGAACTACACGCAGCGC	CTCATTTTG
F244-2 B/C	(119)	AAGGCTGGCATTCTATATAAG	AAGAGAACTACACGCAGCGC	CTCATTTTG
F244-3 B/C	(120)	AAGGCTGGCATTCTATATAAG	AAGAGAACTACACGCAGCGC	CTCATTTTG
F246-2 C	(120)	AAGGCTGGCATTCTATATAAG	AAGAGAACTACACGCAGCGC	CTCATTTTG
F246-3 C	(120)	AAGGCTGGCATTCTATATAAG	AAGAGAACTACACGCAGCGC	CTCATTTTG
F246-4 C	(120)	AAGGCTGGCATTCTATATAAG	AAGAGAACTACACGCAGCGC	CTCATTTTG
F248-1 C	(119)	AAGGCTGGCATTCTATATAAG	AAGAGAACTACACGCAGCGC	CTCATTTTG
F248-2 C	(120)	AAGGCTGGCATTCTATATAAG	AAGAGAACTACACGCAGCGC	CTCATTTTG
F248-4 C	(120)	AAGGCTGGCATTCTATATAAG	AAGAGAACTACACGCAGCGC	CTCATTTTG
F222-5 C	(120)	AAGGCTGGCATTCTATATAAG	AAGAGAACTACACGCAGCGC	CTCATTTTG
Y948-1 C	(120)	AAGGCTGGCATTCTATATAAG	AAGAGAACTACACGGAAGT	GCCTCATTTTG
Y948-2 C	(120)	AAGGCTGGCATTCTATATAAG	AAGAGAACTACACGGAAGT	GCCTCATTTTG
Y949-1 C	(120)	AAGGCTGGCATTCTATATAAG	AAGAGAACTACACGCAGCGC	CTCATTTTG
Y949-2 C	(119)	AAGGCTGGCATTCTATATAAG	AAGAGAACTACACGCAGCGC	CTCATTTTG
Y951-1 C	(119)	AAGGCTGGCATTCTATATAAG	AAGAGAACTACACGCAGCGC	CTCATTTTG
Y951-2 C	(119)	AAGGCTGGCATTCTATATAAG	AAGAGAACTACACGCAGCGC	CTCATTTTG
Y967-1 C	(119)	AAGGCTGGCATTCTATATAAG	AAGAGAACTACACGCAGCGC	CTCATTTTG
Y967-2 C	(120)	AAGGCTGGCATTCTATATAAG	AAGAGAACTACACGCAGCGC	CTCATTTTG
Y991-2 C	(120)	AAGGCTGGCATTCTATATAAG	AAGAGAACTACACGCAGCGC	CTCATTTTG
Y991-3 C	(120)	AAGGCTGGCATTCTATATAAG	AAGAGAACTACACGCAGCGC	CTCATTTTG
Y992-1 C	(120)	AAGGCTGGCATTCTATATAAG	AAGAGAACTACACGCAGTGC	CTCATTTTG
Y992-2 C	(120)	AAGGCTGGCATTCTATATAAG	AAGAGAACTACACGCAGTGC	CTCATTTTG
F332-1 B/C	(120)	AAGGCTGGTATTCTATATAAG	AAGAGAACTACACGCAGCGC	CTCATTTTG
F332-2 B/C	(120)	AAGGCTGGTATTCTATATAAG	AAGAGAACTACACGCAGCGC	CTCATTTTG
Y999-1 B/C	(120)	AAGGCTGGCATTCTATATAAG	AAGAGAACTACACGCAGCGC	CTCATTTTG
Y1001-1 B/C	(120)	AAGGCTGGCATTCTATATAAG	AAGAGAACTACACGCAGCGC	CTCATTTTG
Y1001-3 B/C	(120)	AAGGCTGGCATTCTATATAAG	AAGAGAACTACACGCAGCGC	CTCATTTTG
Y994-1 B/C	(120)	AAGGCTGGCATTCTATATAAG	AAGAGAACTACACGCAGCGC	CTCATTTTG
Y994-2 B/C	(120)	AAGGCTGGCATTCTATATAAG	AAGAGAACTACACGCAGCGC	CTCATTTTG
Consensus	(120)	AAGGCTGGCATTCTATATAAG	AAGAGAACTACACGCAGCGC	CTCATTTTG

	170	204
F238-2B	(169)	TGGGTCACCATATTCTTGGGAACAAGAGCTACAGC
F238-3B	(169)	TGGGTCACCATATTCTTGGGAACAAGAGCTACAGC
F244-1B/C	(170)	TGGGTCACCATATTCTTGGGAACAAGAGCTACAGC
F244-2B/C	(169)	TGGGTCACCATATTCTTGGGAACAAGAGCTACAGC
F244-3B/C	(170)	TGGGTCACCATATTCTTGGGAACAAGAGCTACAGC
F246-2C	(170)	TGGGTCACCATATTCTTGGGAACAAGAGCTACAGC
F246-3C	(170)	TGGGTCACCATATTCTTGGGAACAAGAGCTACAGC
F246-4C	(170)	TGGGTCACCATATTCTTGGGAACAAGAGCTACAGC
F248-1C	(169)	TGGGTCACCATATTCTTGGGAACAAGAGCTACAGC
F248-2C	(170)	TGGGTCACCATATTCTTGGGAACAAGAGCTACAGC
F248-4C	(170)	TGGGTCACCATATTCTTGGGAACAAGAGCTACAGC
F222-5C	(170)	TGGGTCACCATATTCTTGGGAACAAGAGCTACAGC
Y948-1C	(170)	TGGGTCACCATATTCTTGGGAACAAGAGCTACAGC
Y948-2C	(170)	TGGGTCACCATATTCTTGGGAACAAGAGCTACAGC
Y949-1C	(170)	TGGGTCACCATATTCTTGGGAACAAGAGCTACAGC
Y949-2C	(169)	TGGGTCACCATATTCTTGGGAACAAGAGCTACAGC
Y951-1C	(169)	TGGGTCACCATATTCTTGGGAACAAGAGCTACAGC
Y951-2C	(169)	TGGGTCACCATATTCTTGGGAACAAGAGCTACAGC
Y967-1C	(169)	TGGGTCACCATATTCTTGGGAACAAGAGCTACAGC
Y967-2C	(170)	TGGGTCACCATATTCTTGGGAACAAGAGCTACAGC
Y991-2C	(170)	TGGGTCACCATATTCTTGGGAACAAGAGCTACAGC
Y991-3C	(170)	TGGGTCACCATATTCTTGGGAACAAGAGCTACAGC
Y992-1C	(170)	TGGGTCACCATATTCTTGGGAACAAGAGCTACAGC
Y992-2C	(170)	TGGGTCACCATATTCTTGGGAACAAGAGCTACAGC
F332-1B/C	(170)	TGGGTCACCATATTCTTGGGAACAAGAGCTACAGC
F332-2B/C	(170)	TGGGTCACCATATTCTTGGGAACAAGAGCTACAGC
Y999-1B/C	(170)	TGGGTCACCATATTCTTGGGAACAAGAGCTACAGC
Y1001-1B/C	(170)	TGGGTCACCATATTCTTGGGAACAAGAGCTACAGC
Y1001-3B/C	(170)	TGGGTCACCATATTCTTGGGAACAAGAGCTACAGC
Y994-1B/C	(170)	TGGGTCACCATATTCTTGGGAACAAGAGCTACAGC
Y994-2B/C	(170)	TGGGTCACCATATTCTTGGGAACAAGAGCTACAGC
Consensus	(170)	TGGGTCACCATATTCTTGGGAACAAGAGCTACAGC

图2 前-前-S基因核酸序列比较. Consensus: 比较后推断的一致性序列.

2.3 前-前-S基因蛋白质序列分析 全部31株克隆编码的氨基酸序列比较见图3. 前-前-S多肽全长45 aa, 总一致率为75.6%, 可分为2部分, 1-20 aa为高变区, 总一致率为60.0%; 21-45为高保守区, 总一致率为88.0%. 由于图2展示出核苷酸序列中存在的个体化突变现象, 相应的氨基酸序列也存在个体化突变现象. 对应于核苷酸序列的变化, F244患者的3株克隆在第7位氨基酸残基表现出S→L突变(箭头前氨基酸代表其他14个患者的编码序列, 箭头后氨基酸代表某患者的特异性变异, 下同), F332患者2个克隆在第20位氨基酸残基表现出A→V突变, Y948患者2个克隆在第29位氨基酸残基表现出A→E突变, Y1001患者2个克隆在第33位氨基酸残基表现出F→S突变, 符合核苷酸序列中表现出的个体化突变现象.

f248-2	(1)	MQFIITSKLGIYILCGRRAFYIREKLHAAPHFVGHILGNKSYS
f248-4	(1)	MQLIITSKLGIYILCGRRAFYIREKLHAAPHFVGHILGNKSYS
f332-1	(1)	MQLIITSKPGIYILCGRLVFYIREKLHAAPHFVGHILGNKSYS
f332-2	(1)	MQLIITSKPGIYILCGRLVFYIREKLHAAPHFVGHILGNKSYS
y948-1	(1)	MQLIITSKLGIYILCGRLAFYIREKLHEVPHFVGHILGNKSYS
y948-2	(1)	MQLIITSKLGIYILCGRLAFYIREKLHEVPHFVGHILGNKSYS
y949-1	(1)	MQLIITSKLGIYILCGRRAFYIREKLHAAPHFVGHILGNKSYS
y949-2	(1)	MQLIITSKLGIYILCGRRAFYIREKLHAAPHFVGHILGNKSYS
y951-1	(1)	MQLIITSKLGIYILCGRLAFYIREKLHAAPHFVGHILGNKSYS
y951-2	(1)	MQLIITSKLGIYILCGRLAFYIREKLHAAPHFVGHILGNKSYS
y967-1	(1)	MQLIITSKLGIYILCGRLAFYIREKLHAAPHFVGHILGNKSYS
y967-2	(1)	MQLIITSKLGIYILCGRLAFYIREKLHAAPHFVGHILGNKSYS
y991-2	(1)	MQLIITSKPGIYILCGRLAFYIREKLHAAPHFVGHILGNKSYS
y991-3	(1)	MQLIITSQPGIYILCGRLAFYIREKLHAAPHFVGHILGNKSYS
y992-1	(1)	MQLIITSKLGIYILCGRLAFYIREKLHAVPHFVGHILGNKSYS
y992-2	(1)	MQLIITSKLGIYILCGRLAFYIREKLHAVPHFVGHILGNKSYS
y994-1	(1)	MQLIITSKLGIYILCGRLAFYIREKLHAAPHFVGHILGNKSYS
y994-2	(1)	MQLIITSKLGIYILCGRLAFYIREKLHAAPHFVGHILGNKSYS
y999-1	(1)	MQLITITSLGIYILCGRRAFYIREKLHAAPHFVGHILGNKSYS
y1001-1	(1)	MQLIITSKLGIYILCGRLAFYIREKLHAAPHSVGHILGNKSYS
y1001-3	(1)	MQLIITSKLGIYILCGRLAFYIREKLHAAPHSVGHILGNKSYS
Consensus	(1)	MQLIITSKLGIYILCGRLAFYIREKLHAAPHFVGHILGNKSYS

图3 前-前-S多肽氨基酸序列比较. Consensus: 比较后推断的一致性序列

### 3 讨论

本研究组在近期的研究中着重阐述了HBV准种现象<sup>[17-24]</sup>, 在此基础上, 我们发现以往界定的HBV前-S1基因之前存在一个ORF<sup>[25]</sup>, 将其命名为前-前-S基因. 为了证实前-前-S基因的实际存在, 采取横断面研究方式对慢性乙型肝炎患者体内感染的HBV基因组是否存在前-前-S基因进行了研究.

本研究随机选择了HBV DNA阳性的患者, 首先进行了HBV基因型分型研究. Naito et al<sup>[13]</sup>建立的快速分型鉴定方法是基于不同HBV基因型之间存在一定的序列差异, 研究者经过比较后选择相对较为特异的区域设计为一端引物, 而在较为保守的区域选择序列设计为另一端引物. 本文中特异性引物为下游引物, 在保守的前-S2中段设计了上游引物. 在进行PCR过程中, 将不同型特异性引物与保守的对应引物配对进行扩增, 只有型特异性引物可以与上游保守引物配合扩增出PCR产物. 经过分型研究, 本研究抽取的15例患者中, 基因型为B型者1例, B/C混合型者5例, C型9例. 由于是随机抽取的样本, 且样本量较少, 上述分型结果并不代表是普遍现象. 进行HBV基因型分型的研究试图阐明前-前-S基因的存在与HBV基因型之间的关系.

自选择的15例患者中将前-前-S基因克隆到T easy载体中, 之后随机选择阳性克隆进行测序, 并将核苷酸序列进行比较. 除患者Y999和F222选择1个克隆进行测序之外, 其他13例患者选择2个以上的克隆测序和序列分析, 一共测序了31个克隆. 经过序列分析发现, 这31例克隆均编码前-前-S多肽, 前-前-S基因ORF与前-S1读码框架一致, 说明无论B型或C型的HBV

f222-5	(1)	MQLIITSKLGIYILCGRLAFYIREKLHAAPHFVGHILGNKSYS
f238-2	(1)	MQLIITSKLGIYILCGRLAFYIREKLHAAPHFVGHILGNKSYS
f238-3	(1)	MQLIITSKLGIYILCGRLAFYIREKLHAAPHFVGHILGNKSYS
f244-1	(1)	MQLIITLKLGIYILCGRLAFYIREKLHAAPHFVGHILGNKSYS
f244-2	(1)	MQLIITLKLGIYILCGRLAFYIREKLHAAPHFVGHILGNKSYS
f244-3	(1)	MQLTITLKLGIYILCGRLAFYIREKLHAAPHFVGHILGNKSYS
f246-2	(1)	MQLIITSKLGIYILCGRRAFYIREKLHAAPHFVGHILGNKSYS
f246-3	(1)	MQLIITSKLGIYILCGRRAFYIREKLHAAPHFVGHILGNKSYS
f246-4	(1)	MQLIITSKLGIYILCGRRAFYIREKLHAAPHFVGHILGNKSYS
f248-1	(1)	MQLIITSKLGIYITLGGRRAFYIREKLHAAPHFVGHILGNKSYS

基因组均编码前-前-S基因。由于扩增时选择的引物距离前-前-S基因的上游或下游较远的地方, 可以较为完整的表现该基因的编码情况。本研究发前-前-S基因编码45 aa, 根据替换突变的频度, 将前-前-S多肽分为2部分, 前-前-S多肽总一致率为75.6%, 其中1-20 aa为高变区, 8个位点发生变异, 总一致率为60.0%; 21-45为高保守区, 仅3个位点发生变异, 总一致率为88.0%。

针对所有31个克隆的分析发现病毒氨基酸序列的变异具有2个特点, 其一是病毒氨基酸序列的变异具有患者个体化突变的特点, 这是近年来我们<sup>[24]</sup>在针对HBV准种研究中的一个重要发现。既往我们<sup>[14, 24]</sup>的研究并未涉及10例患者以上的测序比较分析, 提出HBV变异具有个体化特点的假说, 而本文重点研究了15例患者以及31个克隆, 经过详细的序列分析研究, 发现在一定的特殊位点会出现相对特异性的变异, 这种变异以替换突变为主, 具有较为强烈的个体化特征, 即突变表现为来自某患者的所有克隆的编码氨基酸与来自其他患者的HBV编码氨基酸不同, 而来自其他患者的HBV在该位点具有一致性。本文发现: F244患者的3株克隆在第7位氨基酸残基表现出S→L突变, F332患者2个克隆在第20位氨基酸残基表现出A→V突变, Y948患者2个克隆在第29位氨基酸残基表现出A→E突变, Y1001患者2个克隆在第33位氨基酸残基表现出F→S突变, 而其他患者来源的HBV在相应位点表现出相对一致的氨基酸编码特点。这在HBV进化中的意义尚不清楚, 但这种变异的指纹样特点在追踪感染链源头方面可能有意义。变异的另一个特点是在氨基酸序列中某些位点的变异表现出相对的一致性, 比较突出的是第19位氨基酸位点, F246的3个克隆、F248的3个克隆、Y949的2个克隆和Y999的1个克隆编码为R, 而其他11个患者的22个克隆均编码L。L为非极性氨基酸, R为带正电荷的极性氨基酸, 属碱性氨基酸。HBsAg血清型的分型是基于第122位氨基酸的K/R<sup>[26]</sup>和第160位点氨基酸的K/R变异<sup>[27]</sup>, 那么前-前-S基因第19位氨基酸残基的L/R突变是否与全S基因(包括前-前-S、前-S1、前-S2和主蛋白)的血清型分型有关, 还需要进一步研究。

总之, 本研究应用分子流行病学研究方法, 自15例患者的血清中提取了HBV DNA基因组, 之后应用HBV基因型分型技术和PCR-TA克隆-测序技术展示了前-前-S基因的广泛存在, 这些样本来自普通HBV感染者, 结合本研究组针对前-X基因的研究<sup>[8, 11]</sup>, 我们初步推断前-前-S基因的存在是一种普遍现象, 是否编码前-前-S基因与HBV的B或C基因型无关。这个重要发现的意义有待于进一步阐明。

#### 4 参考文献

1 Gelibert F, Mandart E, Fitoussi F, Tiollais P, Charnay P. Nucleotide sequence of the hepatitis B virus genome (subtype ayw)

- cloned in *E.coli*. *Nature* 1979;281:646-650
- 2 Charnay P, Mandart E, Hampe A, Fitoussi F, Tiollais P, Galibert F. Localization on the viral genome and nucleotide sequence of the gene coding for the two major polypeptides of the hepatitis B surface antigen (HBsAg). *Nucleic Acids Res* 1979;7:335-346
- 3 吴祥甫, 周翊钟, 冯宗铭, 李戟平, 夏绍源. 人乙型肝炎病毒-HBV adr血清型基因组的克隆和限制酶切图谱. *中国科学B辑* 1983; 2:162-167
- 4 甘人宝, 储美谨, 沈绿萍, 钱苏雯, 李戟平. 克隆的adr血清型乙型肝炎病毒(pADR-1)的核苷酸顺序. *中国科学B辑* 1986;5:55-65
- 5 成军, 杨守纯. 现代肝炎病毒分子生物学. 第1版, 北京:人民军医出版社, 1997:179-182
- 6 董菁, 李进, 施双双, 皇甫竞坤, 成军, 王勤环, 洪源, 王业东, 李莉, 斯崇文. 乙型肝炎病毒基因组准种与变异特点的研究. *解放军医学杂志* 2002;27:116-118
- 7 董菁, 成军. 乙型肝炎病毒基因组中前-前-S区基因的界定. *世界华人消化杂志* 2003;11:1091-1096
- 8 董菁, 成军. 乙型肝炎病毒基因组中前-X编码基因的界定. *世界华人消化杂志* 2003;11:1097-1101
- 9 董菁, 成军, 杨倩. 乙型肝炎病毒基因组高变区界定的初步研究. *世界华人消化杂志* 2004;12:42-46
- 10 杨倩, 董菁, 成军, 刘妍, 洪源, 王建军, 张树林. 乙型肝炎病毒基因组中前-前-S-编码基因启动子序列的确定及转录活性的鉴定. *解放军医学杂志* 2003; 28:761-762
- 11 杨倩, 董菁, 成军, 刘妍, 洪源, 王建军, 张树林. 乙型肝炎病毒基因组中前-X-编码基因启动子活性的鉴定. *解放军医学杂志* 2003;28:763-765
- 12 中华医学会传染病与寄生虫分会、肝病学会. 病毒性肝炎防治方案. *中华传染病杂志* 2001; 19:56-62
- 13 Naito H, Hayashi S, Abe K. Rapid and specific genotyping system for hepatitis B virus corresponding to six major genotypes by PCR using type-specific primers. *J Clin Microbiol* 2001;39:362-364
- 14 董菁, 成军, 王勤环, 皇甫竞坤, 施双双, 张国庆, 洪源, 李莉, 斯崇文. 乙型肝炎病毒DNA序列异质性及准种特点的研究. *中华医学杂志* 2002;82:81-85
- 15 董菁, 成军, 王勤环, 施双双, 皇甫竞坤, 王刚, 洪源, 李莉, 斯崇文. 乙型肝炎病毒前C/C基因准种与变异特点的研究. *解放军医学杂志* 2002;27:122-123
- 16 董菁, 施双双, 皇甫竞坤, 成军, 王勤环, 李莉, 斯崇文. 乙型肝炎病毒X基因准种特点的研究. *中国病毒学* 2002;17:22-26
- 17 董菁, 施双双, 张国庆, 皇甫竞坤, 洪源, 成军, 王勤环, 李莉, 斯崇文. 乙型肝炎病毒C基因启动子区异质性检测的初步研究. *临床检验杂志* 2002;20:72-74
- 18 董菁, 施双双, 张国庆, 皇甫竞坤, 成军, 王勤环, 李莉. 乙型肝炎病毒表面抗原/抗体同时阳性患者体内S基因序列的分析研究. *中国公共卫生* 2002;18:535-537
- 19 董菁, 成军, 王勤环, 施双双, 洪源, 皇甫竞坤, 王刚, 李莉, 斯崇文. 乙型肝炎病毒逆转录酶区基因序列准种与变异研究. *解放军医学杂志* 2001;26:823-825
- 20 皇甫竞坤, 董菁, 邓红, 成军, 施双双, 洪源, 任喜民, 李莉. 乙型肝炎病毒核心基因启动子序列突变及准种. *世界华人消化杂志* 2001;9:1323-1325
- 21 皇甫竞坤, 董菁, 邓红, 成军, 施双双, 洪源, 任喜民, 李莉. 乙型肝炎病毒前S2基因序列异质性的研究. *中华内科杂志* 2002;41:1-4
- 22 成军, 董菁, 刘妍, 李莉, 斯崇文, 王勤环, 张玲霞, 陈菊梅. 乙型肝炎病毒准种研究的临床意义. *世界华人消化杂志* 2002;10:209-221
- 23 董菁, 成军, 皇甫竞坤, 洪源, 王刚, 陈国凤, 李莉, 张玲霞, 陈菊梅. 乙型肝炎病毒序列个体化特征的研究. *解放军医学杂志* 2002; 27:119-121
- 24 董菁, 刘妍, 皇甫竞坤, 施双双, 王刚, 洪源, 陈国凤, 李莉, 陈菊梅, 成军. 乙型肝炎病毒表面抗原基因多态性的初步研究. *胃肠病学和肝病杂志* 2002;11:130-135
- 25 Peterson DL, Paul DA, Lam J, de Bruin TW, Lusic AJ. Antigenic structure of hepatitis B surface antigen: identification of the "d" subtype determinant by chemical modification and use of monoclonal antibodies. *J Immunol* 1984;132:920-927
- 26 Okamoto H, Tsuda F, Sakugawa H, Sastrosoewignjo RI, Imai M, Miyakawa Y, Mayumi M. Typing hepatitis B virus by homology in nucleotide sequence: comparison of surface antigen subtype. *J Gen Virol* 1988;69:2575-2583
- 27 杨倩, 刘妍, 成军, 王建军, 洪源, 张树林. 乙型肝炎病毒X蛋白反式激活作用的研究. *世界华人消化杂志* 2003;11:1248-1230