

# 广东麻鸭 DHBV 全基因克隆及新读码框的发现

何金洋<sup>1\*\*</sup>, 葛文华<sup>2</sup>, 朱宇同<sup>1</sup>, 郭兴伯<sup>1</sup>, 杨瑞仪<sup>1</sup>, 冯丽玲<sup>1</sup>, 张奉学<sup>1</sup>, 陈鸿珊<sup>3</sup>

(1. 广州中医药大学热带医学研究所, 广东广州 510405; 2. 广州中医药大学温病学教研室, 广东广州 510405; 3. 中国医学科学院医学生物技术研究所, 北京 100050)

## Cloning of DHBV Complete Genome from Guangdong Brown Ducks and Finding of a New ORF

HE Jin-yang<sup>1\*\*</sup>, GE Wen-hua<sup>2</sup>, ZHU Yu-tong<sup>1</sup>, GUO Xing-bo<sup>1</sup>, YANG Rui-yi<sup>1</sup>,  
FENG Li-ling<sup>1</sup>, ZHANG Feng-xue<sup>1</sup>, CHEN Hong-shan<sup>3</sup>

(1. Tropical Medicine Institute of Guangzhou University of Traditional Chinese Medicine, Guangzhou 510405, China; 2. Department of Warm Diseases of Guangzhou University of Traditional Chinese Medicine, Guangzhou 510405, China; 3. Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100050, China)

**Abstract:** 3 samples of *Duck hepatitis B virus* (DHBV) positive serum from Guangdong brown ducks were chosen to be experimental samples. One pair of primers named Q1 and Q2 were designed to amplify the DHBV complete genome. Following the DHBV complete genomes from the 3 DHBV positive serum samples were amplified and cloned, these DHBV complete genomes were sequenced and then analyzed through related software and related molecular biology methods. The results showed that all the 5 complete genomes from Guangdong brown ducks were 3027 nucleotides long. After submitting to the GenBank, the accession Number were attained as: A Y433937, A Y521226, A Y521227 (from DHBV positive serum Number 1.); A Y392760, A Y536371 (from DHBV positive serum Number 2 and 3 respectively). One single nucleotide mutation was occurred in the PreS/ S ORF of A Y521227. A new ORF was found in the upper reaches of PreC/ C ORF and named as HORF temporarily. The HORF also exists in the other 8 DHBV complete genomes from Genbank. Phylogenetic analyze indicated that the differentiation degree of the DHBV complete sequences from Guangdong brown ducks were highest in all the DHBV complete sequences exist in the Genbank presently. The successful clone and analysis of DHBV complete genome from Guangdong brown ducks provide further studies with helpful information.

**Key words:** *Duck hepatitis B virus* (DHBV); Complete genome sequence; Clone; Open reading frame.

**摘要:** 选取 3 份 DHBV 阳性广东麻鸭血清提取 DHBV DNA, 设计一对扩增 DHBV 全基因序列引物 Q1 和 Q2, 扩增并克隆 DHBV 全基因序列, 测序并运用相关计算机软件及方法分析获得的全基因序列。结果表明, 广东麻鸭 DHBV 全长为 3027bp。提交 GenBank 后获得的收录号分别为: A Y433937、A Y521226、A Y521227 (来自 1 号血清); A Y392760、A Y536371 (分别来自 2、3 号血清)。A Y521227 的 PreS/ S ORF 出现了单碱基突变, 在 PreC/ CORF 之前发现一个新的 ORF, 暂命名为 HORF, HORF 也同时存在于 GenBank 中储存的另外 8 个 DHBV 全基因序列中。系统发育分析表明, 广东麻鸭 DHBV 在分化程度上是目前储存于 GenBank 中的 DHBV 全序列中最高的。成功克隆了广东麻鸭 DHBV 全基因序列, 序列分析为进一步研究广东麻鸭 DHBV 提供了有益的信息。

**关键词:** 鸭乙型肝炎病毒 (DHBV); 全基因序列; 克隆; 读码框

收稿日期: 2004-12-20, 修回日期: 2005-01-18

\*\* 通讯作者: 何金洋 (1975 - ), 男, 博士, 助理研究员, 主要从事病毒学及中草药抗病毒研究。

Corresponding author. Tel: 020-31774402, E-mail: sunny12345678\_89@yahoo.com.cn

中图分类号: S852.65

文献标识码: A

文章编号: 1003-5125(2005)04-0399-05

目前全世界大约有 4 亿人感染了乙肝病毒 (*Hepatitis B virus*, HBV)<sup>[1]</sup>, 每年因此有 30 万人成为肝癌患者, 并有相似数量致命的胃肠道大出血和腹水患者<sup>[2]</sup>; 每年因 HBV 感染而死亡的人数为 50~120 万<sup>[3]</sup>; 全世界有 1/3 的人在其一生中会感染上 HBV<sup>[4]</sup>, HBV 感染率从 2%~20% 不等<sup>[5,6]</sup>, 对于 HBV 的研究一直是热点问题。

鸭乙肝病毒 (*Duck hepatitis B virus*, DHBV) 与 HBV 同属嗜肝 DNA 病毒科, DHBV 与 HBV 在复制形式、致病机理等方面有着众多的相似之处<sup>[7,8]</sup>, 所以鸭乙肝模型一直被用作抗 HBV 药物筛选及 HBV 致病机理研究的动物模型; 对于 DHBV 的研究也揭示了许多 HBV 的特点。目前对于中国鸭来源的 DHBV 的全基因序列研究已有一些报道, 如上海鸭、重庆鸭、安徽鸭 DHBV 等, 发现这些鸭来源的 DHBV 的基因序列各自都有一定的地域特点。本实验室用广东麻鸭乙肝模型进行多年抗乙肝药物研究, 但尚未见广东麻鸭 DHBV 的基因序列特点研究报道。因此我们对广东麻鸭 DHBV 的全基因序列进行克隆, 为广东麻鸭 DHBV 的进一步研究打下良好基础。

## 1 材料和方法

### 1.1 实验材料

LA TaqDNA 聚合酶、dNTP、pMD18-T 载体、分子量 Marker、*EcoRI* 限制性内切酶、IPTG 和 X-gal 均购自 Takara 公司; 血清病毒 DNA 提取试剂盒购自华美生物工程公司; PCR 产物凝胶回收试剂盒购自 Watson 公司; 小量质粒提取试剂盒购自 Omega 公司。受体菌 JM109 由本所杨瑞仪老师惠赠。

### 1.2 血清中 DHBV DNA 提取

DHBV 阳性血清分别取自 3 只先天携带 DHBV 的广东麻鸭 (经斑点杂交证实 DHBV 阳性)。按照血清中病毒 DNA 提取试剂盒操作说明书进行操作, 即将 100 $\mu$ L 血清与 300 $\mu$ L Prolygents<sup>TM</sup> 混匀, 60 $^{\circ}$ C 水浴 15min, 然后加入 400 $\mu$ L 异丙醇, 振荡混匀, 离心弃上清, 再以 40% 异丙醇洗涤, 最后用 50 $\mu$ L 去离子水溶解沉淀并贮存于 -20 $^{\circ}$ C。用一对引物<sup>[9]</sup> P1 和 P2 检测提取物是否含有 DHBV DNA。引物序列为 P1 (sense): 5'-GCGCTTTCCAA GATACTGGA GCCCAA-3' (nt1426-1451); P2 (anti-sense): 5'-CTGGA TGGGCCGTCA GCA GG AT-TA TA-3' (nt2445-2420)。PCR 扩增条件为 94

30s, 55 $^{\circ}$ C (sense) 30s, 72 $^{\circ}$ C 1min, 30 个循环。

### 1.3 PCR 扩增 DHBV DNA 全基因序列

根据 GenBank 中已储存的 DHBV 全基因序列设计一对引物: Q1 (sense)、Q2 (anti-sense), 其序列为: Q1: 5'-ACCCCTCTCTCGAAA GCAA TA-3'; Q2: 5'-GTGTATGTAA GAGCCGTCCAATC -3'。PCR 反应体系用 50 $\mu$ L, 加 5 $\mu$ L 10 $\times$  LA PCR Buffer, 5 $\mu$ L Mg<sup>2+</sup> (25mmol/L), 8 $\mu$ L dNTP (各 2.5mmol/L), 10 $\mu$ L 血清 DHBV 提取物, 0.5 $\mu$ L LA TaqDNA 聚合酶 (5U/ $\mu$ L), 加去离子水至 50 $\mu$ L。反应条件为: 94 $^{\circ}$ C 30s; 52 $^{\circ}$ C 1min; 72 $^{\circ}$ C 3.5min, 30 个循环, 最后 72 $^{\circ}$ C 延伸 5min。PCR 产物用凝胶回收试剂盒纯化后再用 1% 琼脂糖凝胶电泳鉴定。将 1 号血清扩增 3 份, 2、3 号血清各扩增 1 份。共扩增 5 份 DHBV 全序列。

### 1.4 分子克隆与序列分析

参照常规的氯化钙法制备 JB109 感受态细胞<sup>[10-12]</sup>。分别取上述纯化后的 DHBV PCR 产物 4 $\mu$ L, 与 pMD18-T 载体相连, 然后转化入 JM109 感受态细胞中, 经蓝白斑筛选获得可能插入目的 DNA 的白色菌落, 培养后用少量质粒提取试剂盒提取质粒, 酶切鉴定后送博雅公司进行测序, 测序结果用相关软件进行分析。

## 2 结果与分析

### 2.1 DHBV DNA 提取

三种血清经过 DHBV DNA 提取并经 PCR 扩增后, 其产物经 1.2% 琼脂糖凝胶电泳, 均可见扩增出约 1.0 Kb 的片断, 与预期相符。

### 2.2 PCR 扩增 DHBV 全基因序列

按照前述方法扩增 DHBV 全基因序列 DNA, 经 1% 琼脂糖凝胶电泳, 得到约 3kb 的条带 (见图 1)。

### 2.3 DHBV 全基因克隆

PCR 扩增 DHBV 全基因序列产物经凝胶回收纯化, 与 pMD18-T 载体相连接, 转化入 JM109 感受态细胞, 经蓝白斑筛选, 1 号血清 DHBV 扩增克隆产物挑出 6 个白色细菌, 2、3 号血清 DHBV 扩增克隆产物各挑出 2 个白色细菌, 培养后, 进行少量质粒提取, 然后用 *EcoRI* 酶切鉴定, 均得到预期结果 (见图 1); 至此可以初步确定, DHBV 全基因序列已被成功克隆。

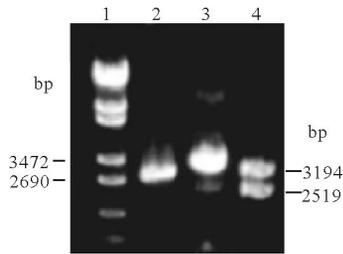


图1 DHBV 全基因扩增与酶切鉴定

Fig. 1 Analysis of PCR product and enzyme cutting of DHBV complete genome

Lane1, -EcoT14I digest molecular weight marker; 2, PCR product of DHBV complete genome; 3, Cloning product of DHBV complete genome; 4, EcoRI enzyme cutting of DHBV complete genome cloning product.

## 2.4 广东麻鸭 DHBV 全基因测序结果

分别在来源于 1 号血清的 3 份重组体和 2、3 号血清的各 1 份重组体中各挑选出 1 个重组体,制成穿刺菌保,送交测序,测序结果整理分析后提交 Genbank,新添编号为:A Y392760、A Y536371、A Y433937、A Y521226、A Y521227。

## 2.5 广东麻鸭 DHBV 全基因序列分析

### 2.5.1 基因结构分析

所得 5 个 DHBV 基因序列全长均为 3027bp,用 ORFfinder 软件分析后得到一般 DHBV 均具有的读码框(open reading frame, ORF)为:P-ORF(20-2536nt),preS/S-ORF(除 A Y521227 为 693-1805nt 外,其余 4 个序列均为 693-1793nt),preC/C-ORF(2524-3027-414nt);A Y521227 的 preS/S-ORF 的终止密码出现了单碱基突变,即 T1792C 的突变,导致终止密码有 nt1791-nt1793 的 TAG 移码突变为 nt1803-nt1805 的 TAA。经过重复测序证明该突变确实存在。该密码的突变到底有何意义?首先使 preS/S 蛋白的末端多出 4 个氨基酸:WEYG,其中 W 为有极性的氨基酸,E 为带电荷的氨基酸,Y 为疏水的氨基酸,G 为甘氨酸。用 SOSUI 软件分析其疏水性平均值为:0.819298,该蛋白有 3 个跨膜螺旋,分别位于:1 8-29aa-ILAGLIGLLVSFLLIKILEIL;2 80-102aa-FIIFLLILLVAA GLLYLTDNMSI;3 127-149aa-QKSLVALMFGLLLIWMTSSSA TQ;如果人为去掉 A Y521227 的 S 蛋白终止密码突变后,预测得到的疏水性平均值为 0.875449,其跨膜区不变,说明疏水性在 S 蛋白终止密码突变后稍有增加,但该突变对 S 蛋白的抗原性应没有多少影响。

另外,在这 5 株 DHBV 中,都发现了一个比较小的阅读框,其起始密码均位于 nt2305,在来源于 1 号血清的 DHBV 全序列(A Y433937、A Y521226、

A Y521227)中,其终止密码位于 nt2424;在来源于 2、3 号血清的 DHBV 全序列(A Y392760、A Y536371)中,其终止密码均位于 nt2430。该读码框与 preC/C 处于同一编码框架,在该读框的-6~-11nt 处为 ATTAAA 的 TA 富集区,提示存在新的启动子的可能性;在对 GenBank 中所储存的所有 DHBV 全序列进行读码框分析后发现除这 5 株广东麻鸭 DHBV 外,另有 8 株 DHBV 存在该读码框,更表明了该读码框存在的可能性,暂命名为 HORF。在这 8 株 DHBV 中,有 6 株为中国鸭来源的 DHBV,且这 6 株也是除 5 株广东麻鸭外所有的在 GenBank 中来源明确的中国鸭来源的 DHBV,另外 2 株分别来源于澳大利亚和德国。也就是说,所有明确来源于中国鸭的 DHBV 均有 HORF,而来源于国外鸭的 DHBV 绝大部分无此 ORF,因此,可以认为,具有 HORF 可看作是来源于中国鸭的 DHBV 的一个显著特点。HORF 若依据起止于 2305-2430nt,则其预测得到的氨基酸序列为:MYTFRFYFCLPQTLSDVTMAFCCASQTTAISYTSVLCCKQIZ。分析其疏水性表明,H 框所预测的编码的多肽是一个可溶性多肽,计算得其疏水性平均值为 0.563415。

在这 5 个广东麻鸭 DHBV 全基因序列中,C 基因启动子均位于 2518-2536nt 内,在 C 基因转录起始点之前有 TATA 盒,在 2423-2431nt 和 2503-2512nt 也存在类似 TATA 盒,正好位于 PreC 转录起始点的上游。DHBV 复制的关键部位 DR1 盒 DR2 分别位于 2541-2552nt 和 2483-2494nt,其序列为 TACACCCCTCTC。其多聚腺苷酸结合位点位于 2778-2783nt,各转录产物均在此处终止。增强子位于 2175-2354nt,增强子保守序列 AGTGTTC T 位于 2221-2230nt。GATA 家族转录因子结合序列 GATA 也出现在此增强子中。这 5 株广东麻鸭的 S 区位于 2566-2622nt,5E 区位于 2346-2386nt,M 区位于 800-840nt,这些都是与 DHBV 的复制密切相关的区段。

### 2.5.2 蛋白质分析

在这 5 个 DHBV 全序列中,C 蛋白翻译均始于 2653nt,含有 262 个氨基酸,4-6aa 是糖基化位点 NAS,221-229aa 为 PreC 蛋白 C 末端加工信号序列 TTVVYGRRR。P 蛋白的 RT 活性部位 YMDD 位于 513-516aa。与其它 DHBV 一样,广东麻鸭 DHBV 的 P 蛋白第 96 位氨基酸亦为酪氨酸,通过此酪氨酸残基,P 蛋白与 dGTP 共价结合,引发副链 DNA 的合成。PreS 蛋白合成起始于 PreS/S-ORF 中第 801ntAUG,含有 330 个氨基酸。S 蛋白中 48、

65、69 位的半胱氨酸和 60 位的组氨酸都是保守的, 这些氨基酸残基与亚病毒颗粒分泌有关。S 蛋白中糖基化信号 Asn-X-Ser/Thr 在所有已报道的 DHBV 中都很保守, 广东麻鸭也不例外, Asn-Gly-Ser 出现在 99-101aa。

### 2.5.3 多序列对比和系统发育分析

用 Blast 软件对比分析发现, 在这 5 株广东麻鸭中, 来自 1 号血清的 DHBV 全核苷酸序列 (A Y433937、A Y521226、A Y521227) 相互之间的核苷酸水平上的同源性为 98%, 而来源于 2、3 号血清的 DHBV 全序列 (A Y392760、A Y536371) 之间的同源性亦为 98%。若将 3 个来自 1 号血清的 DHBV 序列 (A Y433937、A Y521226、A Y521227) 与分别来源于 2、3 号血清的 DHBV 序列 (A Y392760、A Y536371) 之间进行对比, 则其同源性为 92%。说明在广东麻鸭同一种群中的 DHBV 全基因序列之间有较大差别。在 Genbank 所储存的 DHBV 序列中, 与 A Y392760 和 A Y536371 同源性最高的是 2 株来源于德国鸭的 DHBV, 编号为: X58568 和 X58569, 其同源性为 93%。与 3 个来源于同一血清的 DHBV 全基因序列 (A Y433937、A Y521226、A Y521227) 同源性最高者为来自于上海麻鸭的 DHBV, 编号为 M32990, 同源性为 96%。

用目前 GenBank 中存储的所有 DHBV 全基因序列构建了进化树以表明广东麻鸭与其它国家和地区来源的 DHBV 的进化关系 (图 2)。

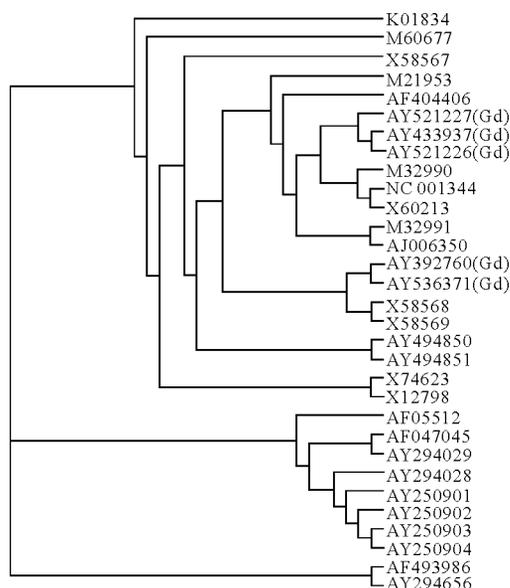


图 2 GenBank 中 DHBV 全基因序列的进化树

Fig. 2 The phylogenetic tree DHBV complete genomes in the GenBank

Gd: form Guangdong brown duck.

从以上进化树可以看出, GenBank 中储存的所有 DHBV 全基因序列在核苷酸水平上的相似性可被分为 3 个大组, 其中所有来源于中国鸭的 DHBV 在同一大组, 所有来源于南非鸭的 DHBV 在同一大组, 来源于澳大利亚、加拿大、美国加利福尼亚鸭的 DHBV 与南非鸭 DHBV 在同一大组; 说明在基因序列上, DHBV 确实存在着明显的地域相似性。3 个来源于 1 号血清的广东麻鸭 (A Y433937、A Y521226、A Y521227) 序列与上海麻鸭 DHBV 及另外 2 株中国鸭 DHBV 在同一小组, 而来源于 2、3 号血清的广东麻鸭 DHBV (A Y392760、A Y536371) 则与 2 株德国鸭 DHBV 在同一小组, 说明其在基因序列上亲缘关系较近。另有 1 大组只有 2 个序列, 1 株来源于印度鸭 DHBV, 另一株来源不祥, 但由广东南方医院传染科提交。

## 3 讨论

曾有学者在 HBV 负链开环处的 DR1 重复序列设计一对引物覆盖了 DR1, 从而用 LAPCR 方法高效率的扩增了 HBV 全基因序列, 并提出扩增 HBV 全基因组的主要难点在 DR1 处<sup>[14]</sup>。我们在摸索扩增 DHBV 全基因序列的 PCR 方法时, 确实发现当 PCR 目标产物跨越了正、负链缺口 (DR1 和 DR2) 时, 稍长一些的 (1kb) 目标片段很难被扩增, 即使较短的目标产物 (1kb 以下), 其扩增效率也大大降低。而当参照了 HBV 全序列扩增方法, 以覆盖 DR1 的引物扩增 DHBV 时, 经过数对引物实验和多种条件优化, 最终并未得到目标 DHBV 全基因序列。后来我们以覆盖 DR2 的一对引物进行实验, 结果很容易得到了 DHBV 全基因序列扩增产物。但当我们用覆盖 HBV 的 DR2 的一对引物扩增 HBV 全基因组时, 最终并未成功。为何 DHBV 与 HBV 需要用不同的引物设计原则才能较好的扩增其全基因组? 这可能反映了 DHBV 与 HBV 基因结构的差别。

HORF 位于 preC/ CORF 之前, 其位置与 HBV 中的 XORF 相似, 但与 Chang 等<sup>[13]</sup>发现的 DHBV-XORF 有较大差别。HORF 是否在体内表达? 有待于进一步研究。经检索, 预测的 HORF 表达的多肽与低等生物细菌及高等哺乳类动物的某些氨基酸有一定同源性, 因此对于 HORF 的进一步研究可能在病毒的进化方面有意义。

3 个来源于广东麻鸭 1 号血清的 DHBV 全基因序列 (A Y433937、A Y521226、A Y521227) 与来源于 2、3 号血清的广东麻鸭 DHBV 全序列均来源于

广东,且出自同一鸭群,为何差别较大?说明虽然在同一地域,甚至同一鸭群,DHBV在进化上却可以是相对独立的。分子进化的始动因素在于基因突变,DHBV复制循环中有逆转录过程,会出现较多的突变,如来源于同一只鸭体内的3个LRDHBV核苷酸的相似性仅为98%,氨基酸水平上的相似性更低,所以DHBV的进化较快。从进化树的总体来看,DHBV总共分为3个大的种群,中国鸭DHBV所在的种群的基因序列分化程度最高,分支最多;且中国鸭DHBV都处在进化树靠近末端的树枝上,说明中国鸭在DHBV种群中是高度进化的,5个广东麻鸭DHBV全序列都处在进化树的最末端树枝上,说明广东麻鸭DHBV的进化程度在已测序的DHBV中是最高的。对于广东麻鸭DHBV的研究可能会有助于嗜肝DNA病毒起源的研究。

## 参考文献

- [1] Lin KW, Kirchner J T. Hepatitis B[J]. *Am Fam Physician*, 2004, 69(1): 75-82.
- [2] Lai C L, Ratzu V, Yuen M F, *et al.* Viral hepatitis B[J]. *Lancet*, 2003, 362(9401): 2089-2094.
- [3] Lavanchy D. Hepatitis B virus epidemiology, disease burden, treatment, and current and emerging prevention and control measures[J]. *J Viral Hepat*, 2004, 11(2): 97-107.
- [4] Jung M C, Pape J R. Immunology of Hepatitis B infection[J]. *Lancet Infect Dis*, 2002, 2: 43-50.
- [5] Marjolis H S, Coleman P J, Brown R E, *et al.* Prevention of hepatitis B virus transmission by immunization. An economic analysis current recommendations[J]. *JAMA*, 1995, 274: 1201-1208.
- [6] Kao J H, Chen D S. Global control of hepatitis B virus infection[J]. *Lancet Infect Dis*, 2002, 2: 395-403.
- [7] Summers J, Mason W S. Replication of the genome of a hepatitis B-like virus by reverse transcription of an RNA intermediate[J]. *Cell*, 1982, 20: 403-415.
- [8] Mason W S, Seal G, Summers J. Virus of Pekin ducks with structural and biological relatedness to human hepatitis B virus[J]. *J Virol*, 1980, 36: 829-836.
- [9] Kock J, Schlicht H-J. Analysis of the earliest steps of hepadnaviral replication: genome repair after infectious entry into hepatocytes does not depend on viral polymerase activity[J]. *J Virol*, 1993, 67: 4876-4874.
- [10] Mandel M, Higa A. Calcium-dependent bacteriophage DNA infection[J]. *J Mol Biol*, 1970, 53: 154.
- [11] Cohen S N, Chang A C Y, Hsu L. Nonchromosomal antibiotic resistance in bacteria: Genetic transformation of *Escherichia coli* by R-factor DNA[J]. *Proc Natl Acad Sci USA*, 1972, 69: 2110.
- [12] Sambrook J, Fritsch E F, Maniatis T. *Molecular cloning: A Laboratory Manual* [M]. Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y. 1989.
- [13] Yao E, Schaller H, Tavis J E. The duck hepatitis B virus polymerase and core proteins accumulate in different patterns from their common mRNA[J]. *Virology*, 2003, 311(1): 81-88.
- [14] Guther S, Li B C, Miska S, *et al.* A novel method for efficient amplification of whole hepatitis B virus genomes permits rapid functional analysis and reveals deletion mutants in immunosuppressed patients[J]. *J Virol*, 1995, 69: 5437-5444.