The Functional Study of HCC-associated Mutations on Hepatitis B Virus

LI, Man Shan

A Thesis Submitted in Partial Fulfilment
of the Requirements for the Degree of
Doctor of Philosophy
in
Biochemistry (Medicine)

The Chinese University of Hong Kong
September 2010

Thesis/Assessment Committee

Professor SHAW, Pang-chui (Chair)
Professor TSUI, Kwok-wing (Thesis Supervisor)
Professor WAN, Chi-cheong (Committee Member)
Professor CHOW, Kwok-chong (External Examiner)

Acknowledgement

I would like to express my deepest gratitude to my supervisor, Professor Stephen K. W. Tsui, for his valuable advice and infinite guidance throughout the years of my study. He is the kindest supervisor who always shows me the faith and passion towards scientific research and religion. It is my pleasure to know him, and I am very grateful to him for his continuous encouragement and patience that guide me though the difficulties in my research.

Special thanks are given to Professor Terrence C. K. Lau for his guidance and technical advice on my study. Also to the collaborators, Dr. Henry L. Y. Chan from Department of Medicine and Therapeutics of The Chinese University of Hong Kong (CUHK), Dr. Winnie Yeo from Department of Clinical Oncology, Faculty of Medicine of CUHK, and Dr. Angeline Bartholomeusz from VIDRL for giving me help and advice on clinical aspects.

I would also like to thank my colleagues who have been helpful and supportive. Miss Sophie K. P. Chan and Miss Daisy W. S. Wong for their collaboration as part their study. Dr. Patrick K. S. Ng, Dr. Patrick T. W. Law, Dr. C. H. Wong, Dr. Winnie W. L. Poon for their technical supports, and also to Miss Fion N. Y. Fong, Miss Alice B. Guan, Mr. Jia-Liang Li, Mr. Chor-fung Ng, Miss Sheila S. K. Li, Miss Emily P. Y. Leung, Mr. Wayne J. W. Zhou, Miss Cecilia S. L. Leung and Miss Vivian H. L. Chan for their contribution to a happy and peaceful working environment throughout my years of study.

I would also like to appreciate the examiners of this project for spending time and giving valuable advices on this thesis.

Last but not least, I would like to express my deepest gratitude to my lovely family, especially to my parents, for their infinite support and endless care.

Abstract

Infection of hepatitis B virus (HBV) causes acute and chronic hepatitis and is closely associated with the development of cirrhosis and hepatocellular carcinoma (HCC). Approximately 60-80% of world's HCC is related to HBV, and it is the third most common cause of cancer death in Asia-Pacific region. Almost 400 million people are chronically infected with HBV and one-third was likely to die of complications of cirrhosis, including liver failure and HCC. As there is a shortage of effective curative treatments, detection and prognosis of the risk of cancer development will be essential to improve survival of patients with chronic HBV infection.

A case-control study was previously carried out to identify HCC-associated genomic markers on HBV. Some of them are clustered at the preS1 and X promoter regions of HBV genotype B and core promoter of HBV subgenotype Cs. The functional significance of these markers to the virus was investigated in our study. Our result showed that one of those markers, the G1613A mutation on core promoter, can significantly increase the promoter activity in a genotype-dependent manner and the effect is reversible by the A-to-G back mutation. We have established an *in vitro*

full-length HBV genome transfection system and the result suggested that the G1613A mutation suppressed the e antigen (HBeAg) secretion and enhanced virus DNA production by downregulating the precore (preC) mRNA transcription. In consistence to the clinical study, the mutation was associated to serum HBV DNA level higher than 6 log copies/ml in female HBV carriers in a univariate analysis. In addition, we demonstrated that the G1613A mutation is a hot spot mutation situated on the negative regulatory element (NRE) on the core promoter in an alignment analysis. To further investigate the molecular mechanism of the mutation, two unknown protein complexes had been shown to bind on the NRE. They showed different binding affinity to the G1613-wild-type and A1613-mutant NRE sequence. Moreover, we showed that in vitro synthesized RFX1 protein could bind to the mutated NRE probe at a higher affinity than that to wild-type NRE probe. Overall, our result suggests that the G1613A mutation exerts its effect by differential binding to some proteins via the NRE region. Studying the mechanism of the mutations may provide insights to the viral pathogenesis and HBV-associated HCC, which has long been a health burden in Asia-Pacific countries.

摘 要

感染乙型肝炎病毒（HBV）可以引起急性和慢性肝炎，也是與肝硬化和肝癌的發展密切相關的。全球大約有 60-80% 的肝癌與乙肝炎病毒有關，並且肝癌是在亞太地區第三大最常見的癌症死亡原因。現在全球將近 400 萬人有慢性乙肝病毒感染，其中三分之一可能死於肝硬化併發症，包括肝衰竭和肝癌。由於缺少有效的治療藥物，透過癌症發展的風險檢測和預後評估以提高患者的生存機會對慢性乙肝病毒感染者將是至關重要的。

我們之前進行了一個病例對照研究，以找出與肝癌相關的乙型肝炎病毒的基因組熱點變異。其中的三個基因組熱點變異分別落在乙肝病毒 B 型 preS1 區和 X 區啟動子和乙肝病毒 Cs 型 C 區啟動子。本研究目的是想探討那些位於啟動子上的熱點變異對病毒的功能性意義。結果顯示位於 C 區啟動子的熱點變異 — G1613A 可令啟動子的作用增加，並且對於不同的病毒基因型有不同幅度的影響。當變異了的啟動子，被回復為野生型的時候（A -至- G 的回復突變），啟動子的作用便相對減低了。此外，我們對此熱點變異作更深入的功能分析，已經建立了一個具複製力的全長乙型肝炎病毒基因體外轉移感染系統，結果表明

摘 要

G1613A 突變透過減低 precore mRNA 的生產從而抑制了 e 抗原（HBeAg）的分泌和增強病毒 DNA 的生產。這結果和臨床研究結果是一致的。在單因素分析當中，女性乙肝病毒攜帶者有 G1613A 突變與乙肝病毒於血清的 DNA 水平高於一百萬拷貝 /毫升有明顯的關係。另外，我們在一個對比分析中顯示了 G1613A 突變是位於 C 區啟動子的反式調控元件（NRE）上的一個熱點變異。為了進一步探討基因突變的分子機制，我們發現兩個未知的蛋白複合物可以與 NRE 結合，而且他們對 G1613 野生型和 A1613 突變體的 NRE 序列具有不同的結合親和力。此外，我們發現在體外合成的 RFX1 蛋白結合到突變了的 NRE 探針比野生的 NRE 探針有更高的結合親和力。總括來說，我們的結果表明 G1613A 突變通過某些蛋白複合物於 NRE 結合力的不同來發揮影響。乙型肝炎病毒感染一直是亞太國家的健康負擔，因此研究這個突變的機制有助提供病毒發病機理和乙肝病毒引起肝癌的理解。

Contents

Chapter 4 Discussion

List of Figures

List of Tables

Abbreviations

AMV	Avian myeloblastosis virus
ATP	Adenosine-5'-triphosphate
BCA	Bicinchonic acid
BCP	Basal core promoter
bp	Base pair
C/EBP	CCAAT enhancer binding protein
cccDNA	Covalently closed circular DNA
CMV	Cytomegalovirus
CURS	Core upstream regulatory sequence
DMEM	Dulbecco's Modified Eagle Medium
dNTP	Deoxynucleoside 5'-triphosphate
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
enhII	Enhancer II
FBS	Fetal bovine serum
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
HBcAg	Hepatitis B core antigen
HBeAg	Hepatitis B e antigen
HBs	HBV surface protein
HBsAg	Hepatitis B surface antigen
HBV	Hepatitis B virus
HBx	HBV X protein
HCC	Hepatocellular carcinoma
HNF	Hepatocyte nuclear factor
HRP	Horseradish peroxidase
IgG	Immunoglobulin G
JCBR	Japanese Collection of Research Bioresources Cellbank
LAR II	Luciferase Assay Reagent II
LBA	Luria-Bertani with ampicillin
NEB	New England Biolabs
Nonidet P-40	octyl phenoxylpolyethoxylethanol
NRE	Negative regulatory element
NREBP	NRE-binding protein
nt	nucleotide
ORF	Open reading frame
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
pgRNA	Pregenomic RNA
POL	HBV polymerase
poly A	Polyadenylation
preC	Precore
PVDF	Polyvinylidene fluoride
RFX1	Regulatory factor X1

RPMI	Roswell Park Memorial Institute
S. D.	Standard deviation
S. E. M.	Standard error mean
SDS	Sodium dodecyl sulfate
SDS-PAGE	SDS polyacrylamide gel electrophoresis
SEAP	Secreted alkaline phosphatase
TAE	Tris-acetate-EDTA
TBST	Tris-Buffered saline with Tween 20
TP	Terminal protein
Tris	Tris(hydroxymethyl)aminomethane
URR	Upper regulatory region
VIDRL	Victorian Infectious Diseases Reference Laboratory

Chapter 1 INTRODUCTION

1.1 Health Burden of HBV Infection

1.1.1 HBV infection worldwide

Hepatitis is a general term meaning inflammation of the liver. It can be caused by a variety of different viruses such as hepatitis A, B, C, D and E viruses. Of the many viral causes of human hepatitis, hepatitis B virus (HBV) is of a greater global importance that chronic HBV infection is a considerable burden to health in the Asian-Pacific region (Ganem *et al.*, 2001; Hollinger *et al.*, 2001; Mahoney *et al.*, 1999; Robinson *et al.*, 1995). According to the World Health Organization Report of 2006, about 75% of the world's population live in areas where the rate of chronic HBV infection is higher than 2%, with more than 2,000 million people have been infected by the virus at some time in their lives. Of these, 350-400 million people remain chronically infected and become carriers of the virus (WHO report 2006, Genava, Switzerland).

1.1.2 HBV chronic infection and hepatocellular carcinoma

It is estimated that chronic HBV carriers would have 100 times higher risk to develop HBV-related hepatocellular carcinoma (HCC) compared to uninfected individuals (Beasley *et al.*, 1981). Fig. 1.1 shows the area of high prevalence of HBV infection correlated with the high annual incidence of primary HCC. High endemic areas include south-east Asia and the Pacific Basin (excluding Japan, Australia and New Zealand), sub-Saharan Africa and some countries in Eastern Europe. HBV-related liver diseases results in about 1- 1.5 million deaths each year (Maynard *et al.*, 1988; Maynard *et al.*, 1990). The virus can cause up to 80% of global cases of HCC worldwide, second only to tobacco among known human carcinogen (Bosch *et al.*, 2004; McGlynn *et al.*, 2005; Llovet *et al.*, 2003; Yeung *et al.*, 2005).

HBV is not cytopathic that the host hepatocyte is not killed by HBV infection. Actually, infected heptocyte can survive for many months, releasing large amount of virions and non-infectious particles. The liver becomes damaged is usually due

to the killing of HBV-infected cells by the body's immune system. It was concluded that the immune-system-mediated chronic inflammation of liver, continues cell death and subsequent cell proliferation might increase the frequency of genetic alterations and the risk of cancer (Chisari, 2000; Chisari *et al*., 1985; Ferrari *et al*, 2003; Visvanathan and Lewin, 2006). The permanent liver inflammation resulting in a degeneration and regeneration process confers to the accumulation of critical mutations in the host genome.

The integration of HBV DNA into the host genome occurs at early steps of clonal tumor expansion. Almost all HBV-associated HCCs harbor chromosomally integrated HBV DNA (Beasley *et al*., 1981; Brechot *et al*., 1980; Paterlini-Brechot *et al*., 2003). In many cases, these integrated genomes are characterized by rearrangements and/or partial deletions in the host genome (Thorgeirsson and Grisham, 2002). Integration can cause deregulation of key regulators of cell cycle control, genes regulating cellular signal transduction cascades, proliferation

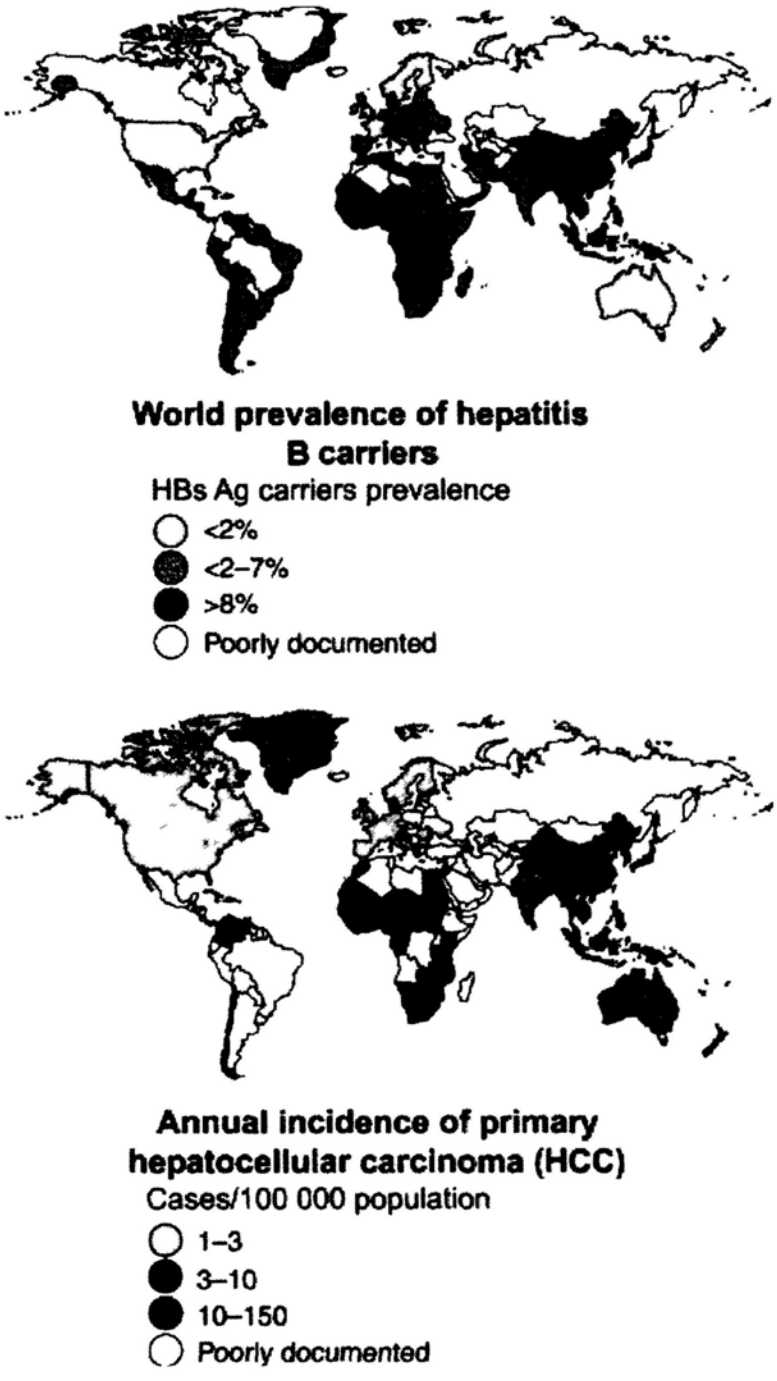
control and cell viability. For example, cyclin A gene (Wang *et al.*, 1990), the retinoic receptor beta gene, the mevalonate kinase gene and sarco/endoplasmic reticulum (ER) calcium ATPase1 gene (Paterlini-Brechot P. *et al.*, 2003, Brechot C. *et al.*, 2000). Recently, hTERT (human telomerase reverse transcriptase) that is part of the telomerase ribonuclear protein complex was found to be targeted in different HBV-associated HCCs (Paterlini-Brechot *et al.*, 2003, Horikawa and Barett, 2001; Hytiroglou, 2006).

Apart from the direct effect of integrated HBV DNA into host genome, HBV X protein (HBx) and PreS2 regulatory protein can induce a tumor promoter-like function. HBx can interferes with signaling cascades that trigger activation of transcription factors like AP-1 (activator protein-1), NF-κB (nuclear factor kappa B), SP1 and oct-1 (Zhang *et al.*, 2006; Waris and Siddiqui, 2003). HBx affects the expression of a variety of genes that are involved in the control of the cell cycle, proliferation or apoptosis. In addition, HBx can interact with p53. It has been

shown that HBx causes a transcription repression of the human p53 gene (Lee and Rho, 2000) and also there is evidence for HBx to bind p53 (Elmore *et al.*, 1997; Ueda *et al.*, 1995). Therefore HBx was considered as a crucial viral protein for the process of HBV-associated carcinogenesis (Cougot *et al.*, 2005; Chan and Sung, 2006; Koike *et al.*, 2002; Staib *et al.*, 2003). Apart from the HBx, the PreS2 protein could also be a trans-activator due to ER stress induced by intracellular retention and accumulation of the protein inside ER. The PreS2 activators bind protein kinase C-α (PKC-α) in the cytoplasm, resulting a DAG (sn-1, 2, diacylglycerol) dependent activation of PKC, which is transduced by the c-Raf-1/MEK/ERK (extracellular signal-regulated kinase) signal transduction cascade (Hildt *et al.*, 2002). This cascade can exert a tumor promoter-like function in carcinogenesis (Boutwell, 1974).

In order to reduce the mortality rate related to HCC in countries with high prevalence of HBV infection, risk stratification of patients is crucial to help

predict the chance of developing HCC after HBV infection. Therefore, the molecular mechanisms underlying HBV-associated HCC development and novel diagnostic markers should be further investigated as therapeutic and preventive targets for HCC.



(Adapted from www.medscape.com/viewarticle/471470)

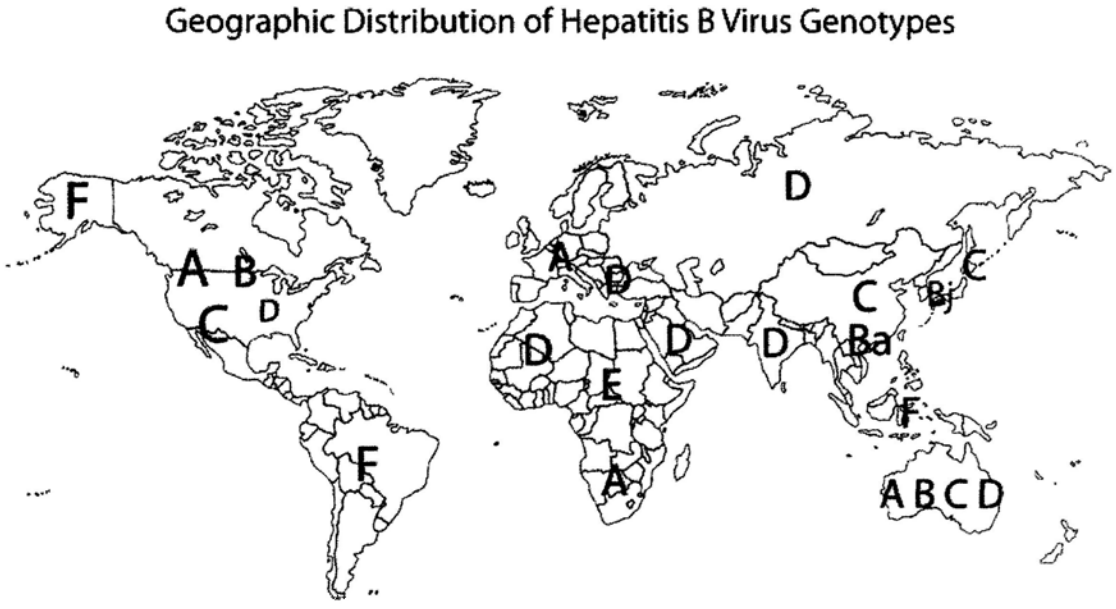
Fig. 1.1 Diagram showing world prevalence of HBV carriers and annual incidence of primary HCC.

1.2 HBV Genotypes and Subgenotypes

HBV has been classified into 8 genotypes (A-H) according to more than 8% divergence in the HBV genome sequence. Different HBV genotypes demonstrate a distinct geographic distribution in the Asian-Pacific region (Chan *et al.*, 2005; Kao *et al.*, 2000; Lesmana *et al.*, 2006; Miyakawa *et al.*, 2003) (Fig. 1.2). HBV genotype A is found predominately in Northern Europe, North America and Central Africa. HBV genotype B is the most prevalent genotype in south-east Asia including the southern China, India, Indonesia and Philippines. HBV genotype C is predominately found in Far East Asia including Japan, Korea, central and northern China. Subgenotypes of HBV genotype B (Bj and Ba) and genotype C (Ce and Cs) have been identified on 4-8% genome divergence. Subgenotype Ce is more prevalent in Far East (Korea and Japan) and northern China whereas Cs is more prevalent in south-east Asia, including Vietnam, Thailand and Malaysia, and southern China. In Hong Kong, HBV genotype C is predominant with 80% is subgenotype Cs (Chan *et al.*, 2005).

Now, more and more studies suggested that HBV genotypes and subgenotypes may differ in clinical manifestations of HBV-related liver diseases and treatment responses. Some studies showed that genotype C is associated with

delayed seroconversion, a more aggressive clinical course, increased risk for HCC and increased resistance to interferon therapy, as compared with genotype B (Chan *et al.*, 2004; Kao *et al.*, 2000). Moreover, subgenotype Ce is associated with increased risk of HCC compared to Cs (Chan *et al.*, 2008). In Japan, subgenotype Ba appears to be associated with more severe liver disease and higher prevalence of HCC then subgenotype Bj (Orito *et al.*, 2001). The actual molecular mechanism leading to the genotypic effect is still unclear, but it has been an evolving area of investigation.



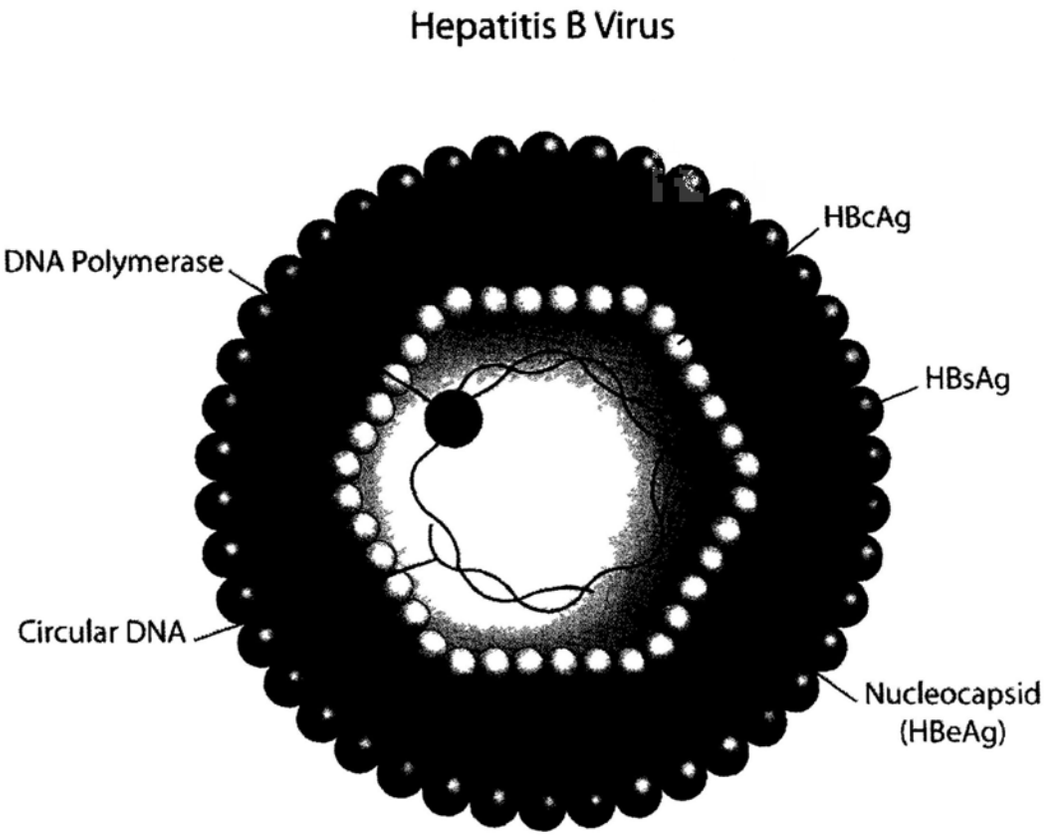
(Adapted from Epidemiology of Infectious Diseases, http://ocw.jhsph.edu)

Fig. 1.2 Geographic distribution of HBV genotypes. Genotypes B and C are prevalent in south-east Asia, including Hong Kong.

1.3 The Hepatitis B Virus

1.3.1 The structure of HBV

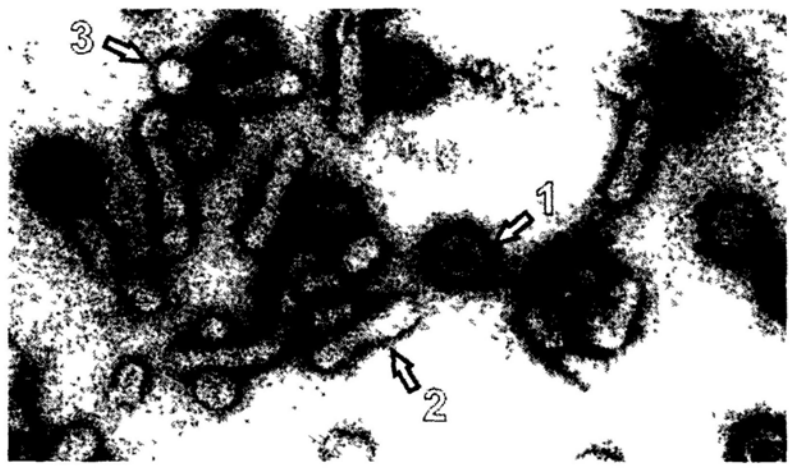
HBV is the smallest hepadnavirus of about 42 nm in diameter, which has a partially double-stranded circular DNA genome of 3.2 kb. The viral genome together with the viral polymerase are packed inside the nucleocapsid core (HBcAg), surrounded by an outer lipoprotein coat (envelope) containing the surface antigen (HBsAg). The schematic diagram of the structure of HBV is depicted in Fig. 1.3. Compared to other human viruses, HBV is unusual in that it contained a vast excess of noninfectious subviral particles composed only of viral surface proteins and lipids, appear predominantly as 22 nm spheres and filaments with a diameter of 22 nm (Fig. 1.4).



(Adapted from Epidemiology of Infectious Diseases, http://ocw.jhsph.edu)

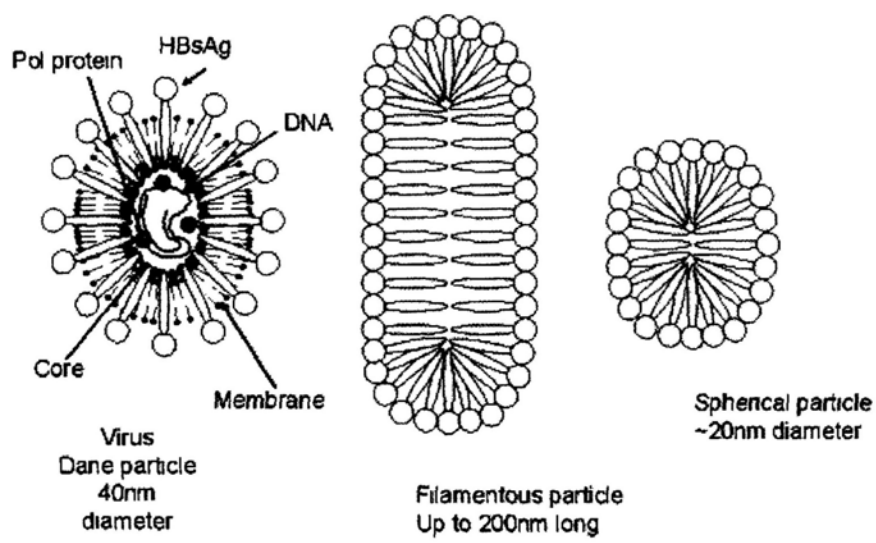
Fig. 1.3 Schematic diagram illustrating the HBV structure. HBcAg: HBV core antigen; HBsAg: HBV surface antigen; HBeAg: HBV e Antigen.

(A)



(Figures adapted from www.klinikum.uni-heidelberg.de)

(B)

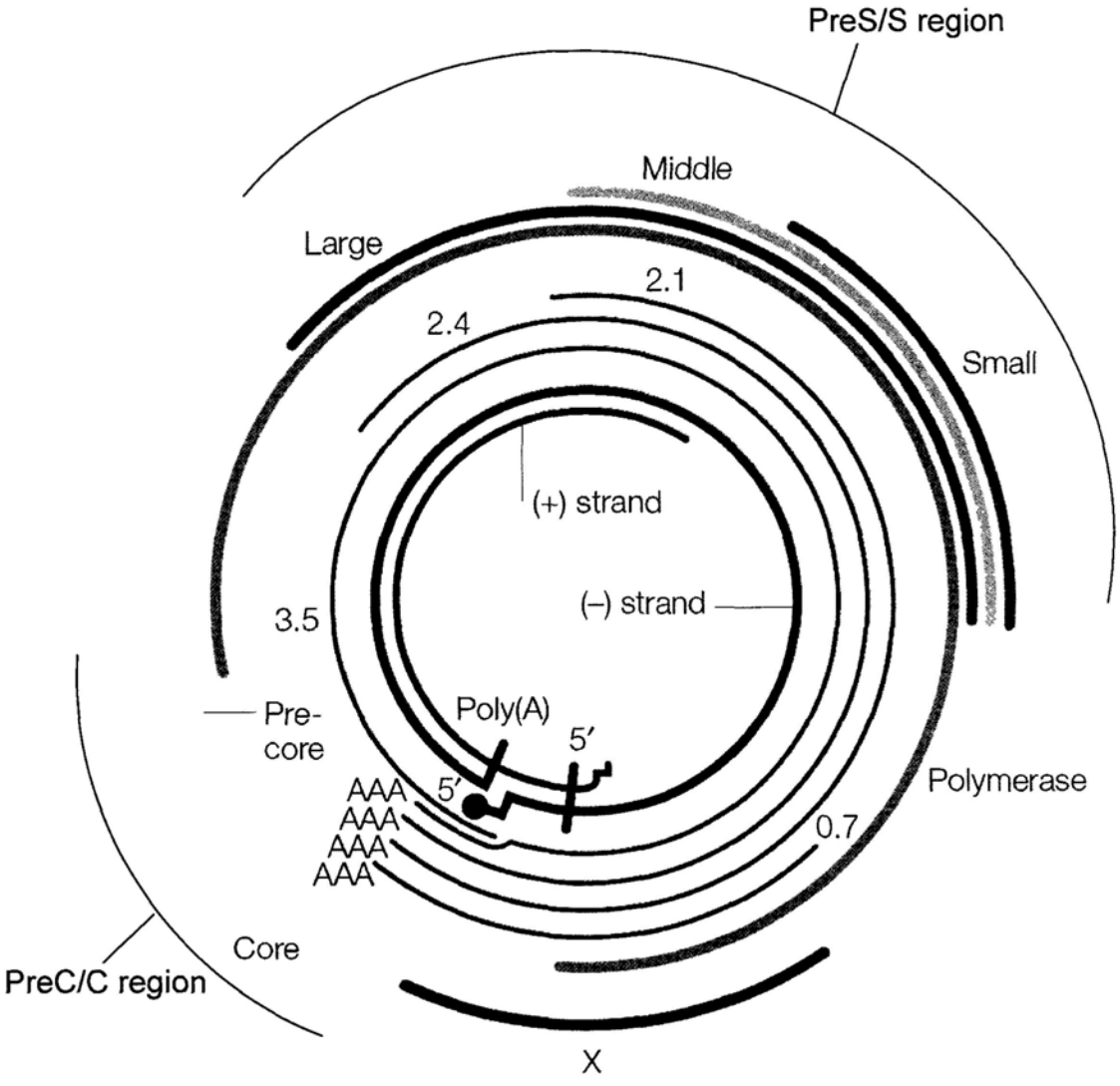


(Adapted from http://pathmicro.med.sc.edu/viral/hepatitis-virus)

Fig. 1.4 An Electron Microscope Presentation of HBV Particles (A) and a schematic diagram showing HBV Dane, filamentous and spherical particles (B). 1: Dane particle; 2: small empty spheres; 3: the filaments; Pol: polymerase.

1.3.2 The genome of HBV

HBV DNA is a relaxed circular, partially double-stranded molecule of 3.2 kb. The positive strand DNA is less than unit length and of variable length (Robinson *et al.*, 1995). The viral polymerase can repair this gap and generate a fully duplex genome. It contains four partially overlapping open-reading frames (ORFs) which codes for seven proteins as shown in Fig. 1.5 as coloured regions. The PreC/C region encodes the core protein and e antigen (HBeAg); the polymerase gene that encodes the viral polymerase-reverse transcriptase; the preS/S region that encodes the large, medium and small surface proteins (HBsAg); and the X gene that encodes a transcriptional transactivator X protein (HBx).



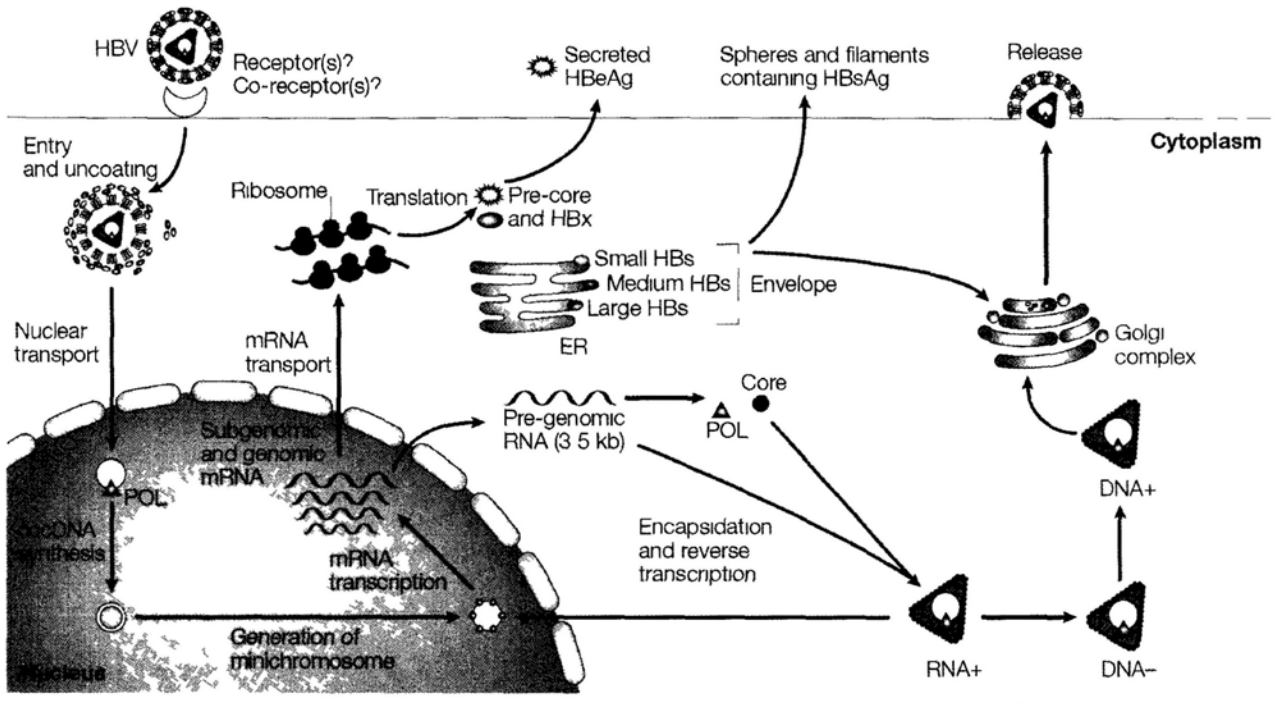
(Adapted from Rehermann and Nascimbeni, 2005)

Fig. 1.5 Schematic diagram showing the HBV genome organization.

1.3.3 The life cycle of HBV

The life cycle of HBV is illustrated in Fig. 1.6. The HBV virion first attach to an unknown receptor on the surface of hepatocytes (Ganem *et al.*, 2004). The viral envelope then fuses with the cell membrane, releasing the nucleocapsid into the cytoplasm. Followed by uncoating, the viral genome and polymerase are transferred into the nucleus, where the semi-circular DNA is repaired by the viral polymerase to yield a covalently closed circular DNA (cccDNA). The cccDNA is the form which HBV persists in host cells and serves as a template for transcription of four viral RNAs. These transcripts are polyadenylated and transported to the cytoplasm. The 3.5 kb pre-core mRNA (preC mRNA) is translated into the precore antigen, which is post-translationally modified to form the e antigen (HBeAg) as a secretory protein. The 2.4 kb mRNA is translated into large, middle and small surface proteins (HBs). The 0.8 kb mRNA is translated into the transcriptional transactivating HBx (Chisari *et al.*, 1997; Ganem and Schneider, 2001; Mahoney *et al.*, 1999; Robinson, 1995). Another 3.5 kb RNA

species, called pregenomic RNA (pgRNA) spans the entire genome, is encapsidated together with the viral polymerase and a protein kinase into the core particles. The pgRNA serves as a template for translation of core protein and viral reverse transcriptase/polymerase (POL). It also serves as a template for the reverse transcription of the negative-strand DNA inside the nucleocapsid. (Ganem and Schneider, 2001; Mahoney *et al.*, 1999). The new, mature viral particles can then either follow the assembly pathway and secreted out, or the genome amplification pathway which the nucleocapsids retained inside the cells as the intracellular reservoir of the cccDNA.



(Adapted from Rehermann and Nascimbeni, 2005)

Fig. 1.6 Figure showing the life cycle of HBV. HBV enters the hepatocytes through an unknown receptor on the cell surface, and the envelope is subsequently removed. Within the nucleus, the partially double-stranded DNA is repaired to form a cccDNA, which serves as the stable template for the transcription of the viral mRNA necessary for productive viral replication. This cccDNA template remains in the nucleus during chronic viral infection and may persist in the liver for the lifetime. The mRNAs are then transported into the cytoplasm and translated into viral proteins. The pgRNA, together with the POL, is encapsidated into the core capsid and serves as a template for the reverse transcription to DNA. Finally, new mature viral particle is released after the assembly of the envelope proteins.

1.3.4 Encapsidation of pgRNA and viral replication

The core protein of HBV can be dimmerized and assembled into nucleocapsids. It is known that the core protein can self-assemble, because capsids can be formed in *Escherichia coli* recombinants expressing the HBV core gene. During the life cycle of the HBV, the pgRNA is packaged into the nucleocapsid in a process called encapsidation, The HBV polymerase-reverse transcriptase binds to the pgRNA within a site, called the *epsilon* (ε) that has high degree of secondary structure. The ε sequence is within the terminal repeat and so is present in both ends of pgRNA. However, the polymerase only binds at the 5′ end. The ε structure with its bound polymerase protein acts as the packaging signal for incorporation of the pgRNA into capsid.

The reverse transcription (RT) domain of viral polymerase (Fig. 1.7) carries out DNA synthesis and the terminal protein (TP) domain acts as the primer for the initiation of the negative strand DNA synthesis. A covalent bond is formed between the –OH group of a tyrosine residue near the N terminus of the

polymerase and the first nucleotide. The pgRNA acts as template for DNA synthesis. Initially a 4-nucleotide DNA of negative strand was synthesized and transferred to a complementary sequence direct repeat 1 (DR1) near the 3′ end of the pgRNA. DNA synthesis continues until the 5′ end of the pgRNA, afterwards the RNase H activity of the polymerase degrades the pgRNA from the DNA-RNA hybrid. All the RNA is removed except for a short sequence of 15-18 bp including the cap, which the –OH group at the 3′ end of the RNA remains as the primer for the positive strand DNA synthesis. The RT and RNase H domain are conserved, and necessary as structural components for pgRNA encapsidation.

At present, no direct structural information on any hepadnaviral polymerase protein is available although homology based models for RT (Das et al., 2001) and RNase H domain (Allen et al., 1998) have been proposed.

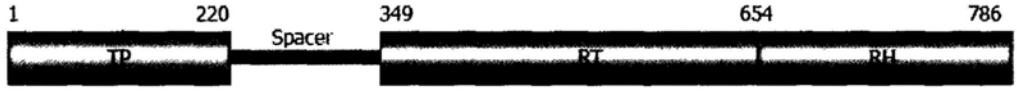


Fig. 1.7. The Schematic diagram of the HBV polymerase. TP: Terminal protein domain; RT: reverse transcriptase domain; RH: RNase H domain.

1.3.5 The mutations on HBV genome

Although mutations can occur randomly along the HBV genome, the overlapping genes of HBV limit the number and location of viable mutants. Mutations in the pre-S/S, preC/C and polymerase regions have been more fully characterized (Ganem *et al.*, 2004; Jean-Michel, 2005; Kao *et al.*, 2003, Kramvis and Kew, 1999; Lesmana *et al.*, 2006; Noboru *et al.*, 2007; Osama and Makoto, 2006). Despite HBV is a DNA virus, its replication is through an RNA-replicative intermediate requiring an active viral reverse transcriptase/polymerase enzyme. However, the enzyme is believed to lack a 3' to 5' exonuclease proofreading capacity (Gunther *et al.*, 1998). The mutation rate of HBV is approximately 2×10^{-4} base substitutions/site/year, which seems to be 10-fold higher than that of other DNA viruses (Gunther *et al.*, 1998). As a result, HBV has a "quasispecies" distribution in infected individuals which implies that HBV circulates as a mixture of genetically distinct but closely related variants that are in equilibrium at a given time point of infection in a given replicative environment. The quasispecies

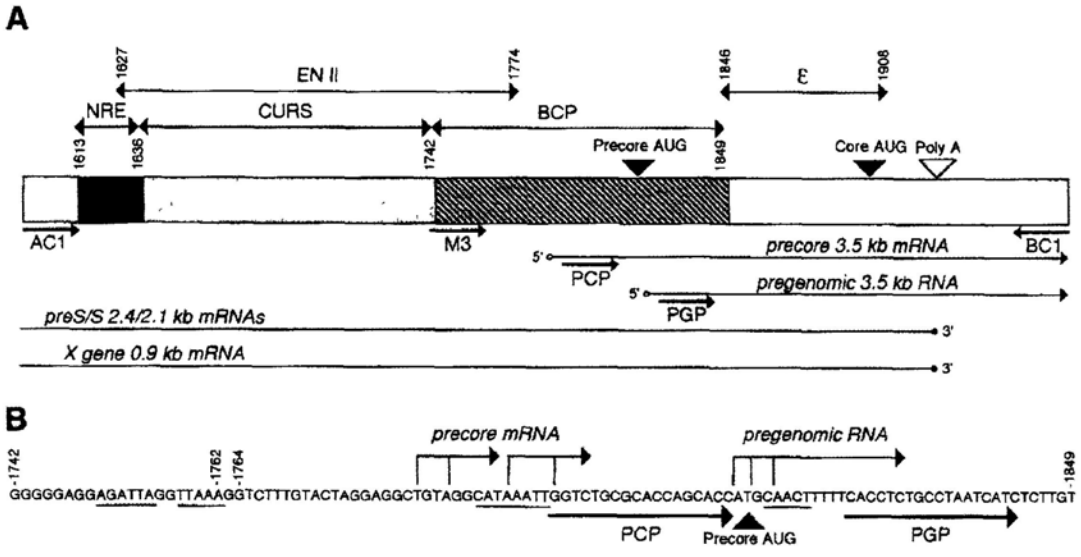
distribution of HBV implies that any newly generated mutations would give a selective advantage to the virus in a given replicative environment that allows the corresponding viral population to overtake the other variants. Under particular selection pressure, both endogenous (host immune clearance) and exogenous (vaccine and antivirals), strongly influence the predominant HBV quasispecies in an infected individual. The fittest virus, i.e. the virus that replicates best, becomes dominant species. Viruses encoding changes associated with antiviral resistance often have reduced replication in vitro, but the accumulation of additional mutations helps restore viral fitness. These compensatory mutations may occur not only in the polymerase gene, but also in other genes such as overlapping envelop gene, the precore gene or in regulatory regions such as the basal core promoter. However, until now, there is no report on mutations on RNaseH domain and terminal domain that can significantly alter the viral replication (Sheldon et al., 2006)

1.3.6 The core promoter of HBV

The HBV core promoter includes 238 bp between nt. 1575 to 1813 in our study (numbering from the unique *Eco*RI site). It overlaps the 3' end of the X ORF and the 5' end of the preC/C region. Fig. 1.8A show the schematic diagram of the structure of core promoter. Basically, the core promoter can be divided into 2 parts, the basal core promoter (BCP) and the upper regulatory region (URR). The BCP (nt. 1742-1849) is defined as the region consisting of *cis*-acting element which initiates the pgRNA and pre-C mRNA transcription (Moolla *et al*., 2002). Multiple start sites of preC mRNA have been identified (Fig. 1.8B), which may due to lack of a canonical TATA box within the BCP (Kramvis and Kew, 1999). The URR consist of the negative regulatory element (NRE, nt. 1613-1636) and the core upstream regulatory sequence (CURS, nt. 1637-1742) overlapping the enhancer II (enhII, nt. 1627-1774). The CURS can activate the BCP in well-differentiated hepatoma cells by 200-2000 folds, in an orientation- and position-dependent manner (Yuh *et al*., 1992). Its stimulatory effect can be abolished by the upstream

NRE. The NRE has a minor inhibitory effect on its own but is strongly repressive in the presence of a functional enhII. Co-operative interaction of various liver-enriched transcription factors is necessary for the liver-specific expression from the core promoter such as HNF-4, HNF-3 and C/EBP (Kramvis and Kew, 1999). Therefore conserved sequence of the core promoter is crucial for maintaining the active viral replication. Any mutations in core promoter may contribute to the persistence of HBV within the host, leading to chronic infection and the risk to HCC. The most frequently reported mutations in the core promoter are the A-to-T and G-to-A mutations at positions nt. 1762 and nt. 1764, respectively (A1762T/G1764A) on the BCP region (named the BCP mutation in this study) (Gunther *et al.*, 1998; Jean-Michel, 2005.). This double mutation is well known to accompany by a reduced level of HBeAg expression in chronic hepatitis and HCC patients (Gunther *et al.*, 1998; Jean-Michel, 2005; Laras *et al.*; 2002), which is believed to exert a selective advantage to the viral propagation.

The HBV core promoter plays a central role in the replication and the life cycle of the virus by directing the synthesis of two types of 3.5-kb viral transcripts that differ at their 5′ initiation sites (Will *et al.,* 1987; Yaginuma *et al.,* 1987). The shorter pgRNA, initiated at nt. 1815 ± 5, serves as the mRNA for the core and polymerase genes and following encapsidation as the template for reverse transcription to generate viral DNA. The slightly larger preC mRNA is initiated 15–35 bp upstream of the pgRNA start sites and contains an additional upstream in-frame translational initiation signal. Translation from this start codon leads to the synthesis of the preC protein, the HBeAg precursor, which is secreted in the serum after posttranslational modification (Takahashi *et al.,* 1991). The synthesis of the preC and pregenomic RNAs is regulated by two separate promoters which the expression can be differentially regulated by ubiquitous (e.g., Sp1) and liver-enriched (e.g., HNF4) transcription factors and a HNF4-binding site, called the DR1 HRE, spanning the TATA box-like sequence of the pre-C promoter (Yu and Mertz, 1996).



(Adapted from Andreas et al., 2002)

Fig. 1.8 Schematic diagram of the HBV core promoter and its transcripts. (A) Schematic representation and organization of the core promoter and precore/core regions of HBV. The negative regulatory element (NRE) nt. 1613–1636, core upstream regulatory sequences (CURS) nt. 1636–1742, basic core promoter (BCP) nt. 1742–1849, and enhancer II (ENII) nt. 1627–1774. The position of the HBV encapsidation signal (nt. 1846–1908) is also indicated. The 5′ ends of the 3.5-kb transcripts and the 3′ ends of the preS/S and X gene transcripts are shown. Closed triangles indicate the translation initiation codons for precore and core and the open triangle the polyadenylation (Poly A) signal for HBV RNA transcripts. (B) Nucleotide sequence of the BCP region (nt 1742–1849). PreC mRNA and pg RNA initiation sites are marked with vertical lines attached to horizontal arrows, indicating the origin and direction of transcription. The TATA boxes and the initiators for preC mRNA and pg RNA are underlined.

1.4 Aims of Our Study

A case-controlled study was previously carried out to find out genomic markers on HBV genome. 100 HBV-related HCC patients and 100 chronic HBV carriers (both age- and sex-matched) were recruited from the Hepatology Clinic in the Prince of Wales Hospital of the Chinese University of Hong Kong. The serum HBV genomes were sequenced. There are 3 identified HCC-related mutations on HBV promoter regions. They are the C1165T mutation on X promoter, T2712C mutation on the preS1 promoter in HBV genotype B, and G1613A mutation on core promoter in HBV subgenotype Cs. In this project, we have several objectives:

I. Construct promoter/luciferase plasmids to study the effect of HCC-related mutations on the promoters in human hepatoma cell lines.

II. Determine the relationship of the mutation(s) and serum viral load in chronic carrier.

III. Prepare the 1.3-fold HBV full-length constructs and determine the effect of the mutation(s) on the *in vitro* viral replication and protein synthesis (HBsAg and HBeAg).

IV. Determine the involvement of regulatory proteins (if any) in regards to the effect of the mutation(s).

In order to reduce the mortality rate related to HCC, and guide resource allocation in Asian countries with high prevalence of HBV infection. risk stratification of patients is important to help predict the chance to develop HCC upon HBV infection. In long term, this study might help in understanding the pathogenesis of HBV-induced HCC development, which has long been a complicated problem in Asia-Pacific countries.

Chapter 2 MATERIALS AND METHODS

2.1 Patient Samples in the Study

2.1.1 Case-control study

One hundred of each HBV-related HCC and HBV-infected non-HCC individuals (carriers) were recruited from the Joint Hepatoma Clinic, Prince of Wales Hospital, Hong Kong from July 1999 to December 2000 (Sung *et al.,* 2008). They were sex- and age-matched. The full-length HBV sequence in their serum were determined by DNA sequencing. Among the 100 cases of HCC, there were 37 cases of genotype B and 63 cases of genotype C (47 cases of subgenotype Cs and 16 cases of subgenotype Ce). In the carriers, 51 cases were infected with genotype B and 49 cases with genotype C (39 cases of subgenotype Cs and 10 cases of subgenotype Ce).

2.1.2 HBV carriers of subgenotyope Cs

Residual serum samples from a cohort of chronic hepatitis B patients undergoing a surveillance program for hepatocellular carcinoma recruited from the Hepatology Clinic, Prince of Wales Hospital from October 1997 to November 2000 was studied (Chan *et al.*, 2008). Based on sequencing of the HBV S gene, 255

chronic carriers were confirmed to have subgenotype Cs HBV infection. The configuration of nucleotide 1613 was determined in the available residual serum samples of these patients by direct DNA sequencing.

2.2 Template Selection for Promoter Study

The HBV sequence with greatest homology to the consensus was chosen as template for polymerase chain reaction (PCR) amplification. The consensus sequence is obtained by aligning multiple patients-derived HBV isolates of the same genotype/subgenotypes and setting the consensus threshold at 80% for each nucleotide. The positions of promoter regions were determined according to Renee McKay (School of Molecular and Biomedical Science, University of Adelaide, Australia) and Benjamin J Dewar (Department of Pathology, the University of Melbourne, Australia).

2.2.1 The X promoter (nt. 980-1373)

Eighty-eight HBV genome sequences from patients' database were aligned with genotype B consensus sequence. HBV isolate each from a carrier and HCC patient, with nucleotide 'C' and 'T' at nucleotide position (nt.) 1165 respectively have 99.8% and 98.2% similarity to the consensus sequence.

2.2.2 The preS1 promoter (nt. 2706-2887)

A HBV isolate from a carrier was chosen as it had 99.9% similarity in pre-S1 promoter region with the consensus sequence of genotype B. Since the pre-S1 promoter sequence of the carrier and HCC only differed at nt. 2712, which was the site of mutation. Therefore only one isolate was chosen for cloning.

2.2.3 The core promoter (nt. 1575-1813)

Two templates were selected for HBV subgenotype Cs, one in carrier and one in HCC. They have 99.9 % similarity with the consensus sequence. They contained the BCP mutation (A1762T/G1764A), with a 'G' and 'A' at nt. 1613 respectively. Only one isolate from carrier was selected for subgenotype Ce and genotype B. They have 92.1% and 99.2% similarity to the corresponding consensus sequence in subgenotype Ce and genotype B respectively. Both isolates contained the wild-type BCP region (A1762/G1764). The HBV DNA sequence between nt. 1587-1823 was missing from one isolates in genotype B (patient number 841). Therefore, there were only 87 HBV sequences for alignment for core promoter region in the template selection for genotype B.

2.3 Construction of HBV Promoter Clones from Patient Isolates

2.3.1 Naturally occurring HBV promoters from patients

A construct of core promoter of genotype A was a generous gift from Dr. Angeline Bartholomeusz of Victorian Infectious Diseases Reference Laboratory (VIDRL), Melbourne. The construct served as a positive control (the BCP mutation) in the luciferase assay. Other core promoter constructs were derived from HBV isolates extracted from patients' serum. The primers used in the amplification of HBV promoters in patients' isolates are listed in Table 2.1-3. The PCR reactions included 0.2 mM dNTP, 0.2 μM each of forward and reverse primers, 5μl of extracted HBV DNA, 0.25μl Advantage^{TM} cDNA polymerase (Clontech Laboratories, Inc.) plus 0.25μl of Pfu DNA polymerase (Promega) and 1X Advantage^{TM} cDNA polymerase PCR buffer (Clontech) in a total volume of 50 μl. The PCR was performed with a 3 min initial denaturation at 94℃, followed by 15 cycles of amplification (94℃ for 36 s, 55℃ for 45 s and 68℃ for 50 s) and a final extension at 68℃ for 10 min. Afterwards, 2 μl of the PCR product was re-amplified in a reaction volume of 50μl with components same as above except for the buffer and DNA polymerase, which were changed to 1X Pfu DNA polymerase buffer (Promega) and 0.5 μl of Pfu DNA

polymerase (Promega). The purified PCR products were cleaved by 0.5 μl each *Xho*I and *Hind*III (Amersham Biosciences) in 1X buffer K, incubated at 37℃ for 4 hours (hr) followed by enzyme inactivation at 85℃ for 15 min. After purification, the digested product was ligated into 2 μl of pGL3-basic vector with T4 DNA ligase (New England Biolabs) in 10X T4 DNA ligase buffer in a total volume of 20 μl and incubated at 4℃ overnight. The ligated product was transformed into competent DH5α cells, and selected on Lurin-Bertani-ampicillin (LBA) agar plate at 37℃ for 16 hr. The confirmed recombinant plasmids were purified by the Rapid Plasmid Purification System (Marligen Biosences) according to manufacturer's instruction. The purified clones were sequenced with vector primers for further confirmation, and prepared in a large-scale by Hi-speed^{TM} Plasmid Midi Kit (Qiagen) according to manufacturer's instruction.

Primer	Isolate	Sequence (5′ to 3′)	Start	End	Direction	Restriction site
B980-XhoI-F	Carrier/ HCC	GGCTCGAGGGAAAG TATGTCAACGAATTG TGGG	980	1004	Forward	*Xho*I
BHX-HINDIII-R	HCC	GCAAGCTTGGGAAT GATGTATACTTGCG GG	1352	1373	Reverse	*Hind*III
BCX-HINDIII-R	Carrier	GCAAGCTTGGAAAT GATGTATACTTGCG GG	1352	1373	Reverse	*Hind*III

Table 2.1 Primers used in cloning HBV X promoters from patient isolates. The restriction enzyme sites were underlined.

Primer	Isolate	Sequence (5′ to 3′)	Start	End	Direction	Restriction site
CBPreS1P-F	Carrier	CCCTCGAGCCAG AGTATGTAGTTAA TC	2706	2724	Forward	*Xho*I
*HBPreS1P-F	HCC	CCCTCGAGCCAG AG**C**ATGTAGTTAA TC	2706	2724	Reverse	*Xho*I
PreS1-R	Carrier/ HCC	CACTAAGCTTTCC CCATGCCTTTTC GAGG	2887	2905	Reverse	*Hind*III

* Primer HBPreS1P-F containing the T2712C mutation was used as the mutagenesis primer, which the site of mutation is indicated as red colour in bold.

Table 2.2 Primers used in cloning HBV preS1 promoters from patient isolates.

HBV Genotype	Primer	Isolate	Sequence (5′ to 3′)	Start	End	Direction	Restriction site
A	ACoreP-R	Given by Dr. Angeline Bartholomeusz	ACTTAAGC TTAGTGCT GGTGCGC AGACC	1813	1796	Reverse	*HindIII*
B	BCoreP-R	Carrier/HCC	ACTTAAGC TTGGTGCT GGTGAAC ACACC	1813	1796	Reverse	*Hind*III
C	*CCoreP-F	Carrier/HCC	CCCTCGA GCCGTGT GCACTTC GCTTC	1575	1592	Forward	*Xho*I
	CCoreP-R	Carrier/HCC	ACTTAAGC TTGGTGCT GGTGAAC AGACC	1813	1796	Reverse	*Hind*III

*Genotype A and B shared the same forward primer as that used in genotype C.

Table 2.3 Primers used in cloning HBV core promoters from patient isolates. The restriction enzyme sites were underlined.

2.3.2 Site-directed mutagenesis of promoter constructs

Mutations were introduced by PCR-based site directed mutagenesis using 2× PicoMaxx® master mix (Stratagene) with appropriate primers (listed in Table 2.4 and 2.5). In the case of preS1 promoter, the cloning primers for the HCC isolates contains the site of mutation. Therefore, no additional mutagenesis primer is needed. In each pGL3-promoter clone, two fragments were amplified. The PCR reaction included 0.2 μM each of forward and reverse primers, 3 μl of purified HBV naturally occurring promoter construct (1:100 dilution) and 25 μl 2× picomaxx mastermix reaction in a total volume of 50 μl. The PCR was performed with a 3 min initial denaturation at 94℃, followed by 32 cycles of amplification (94℃ for 36 s, 55℃ for 45 s and 72℃ for 50 s) and a final extension at 72℃ for 10 min. These 2 fragments are overlapped with a few bases including the target site of mutation. The PCR products were examined in 1% TAE agarose gel and purified by gel extraction kit (Qiagen), followed by amplification without any primers for 10 cycles (94℃ for 36 s, 55℃ for 45 s and 72℃ for 50 s) and a final extension at 72℃ for 10 min. During the extension period, RVprimer3 and GLprimer2 were added and PCR resumed for 20 more cycles. After gel purification, the fragments were digested by *Xho*I and *Hind*III and cloned

into pGL3-basic vector as previously described. The BCP mutation (A1762T/G1764A) was naturally occurring in the core promoter constructs in subgenotype Cs (pC1 and pC2). Construct pC3 and pC4, which contain wild-type BCP (A1762/G1764), were constructed to investigate the effect of G1613A mutation with/without the combination of BCP mutation in subgenotype Cs.

Primer	Isolate	Sequence (5′ to 3′)	Start	End	Mutation	Direction
CX-Mut-F	Carrier	TCGGCAACGG**T**CTGG TCTGTGC	1155	1176	C1165T	Forward
CX-Mut-R	Carrier	GCACAGACCAG**A**CCG TTGCCGA	1155	1176	G1165A	Reverse
BHX-Mut-F	HCC	TCGGCAACGG**C**CTGG TCTGTGC	1155	1176	T1165C	Forward
BHX-Mut-R	HCC	GCACAGACCAG**G**CCG TTGCCGA	1155	1176	A1165G	Reverse

The sites of mutation are indicated as red colour in bold.

Table 2.4 Primers used in site-directed mutagenesis in HBV X promoter.

HBV Genotype/ subgenotype	Primer	Isolate	Sequence (5′ to 3′)	Start	End	Mutation	Direction
B, Cs and Ce	MutCC 3Cor-F	Carrier	**C**GTCGCATGGA**A**A CCACCGTG	1602	1622	G1613A	Forward
	MutCC 3Cor-R	Carrier	**C**ACGGTGGT**T**TCC ATGCGACG	1622	1602	C1613T	Reverse
	MutHC 3Core-F	HCC	**C**GTCGCATGGA**G**A CCACCGTG	1602	1622	A1613G	Forward
	MutHC 3Core-R	HCC	**C**ACGGTGGT**C**TCC ATGCGACG	1622	1602	T1613C	Reverse
Cs	BCP-WT-CC3-F	Carrier	TTGGGGGAGGAG ACTAGATTAA**A**G**G** TCTT	1740	1768	T1762A/ A1764G	Forward
	BCP-WT-CC3-R	Carrier	ACAAAGA**C**C**T**TTA ATCTAGTCTCCTC CCCC	1742	1771	A1762T/ T1764C	Reverse
	BCP-WT-HC3-F1	HCC	TGGGGGAGGAGA CTAGGTTAA**A**G**G**T CTTTG	1741	1770	T1762A/ A1764G	Forward
	BCP-WT-HC3-R1	HCC	CAAAGA**C**C**T**TTAA CCTAGTCTCCTCC CCCA	1770	1741	A1762T/ T1764C	Reverse

(to be continued)

A	MutA-G1613A-F	Clone given by Dr. Angeline Bartholomeusz	CGTTGCATGGA**A**ACCACCGTG	1602	1622	G1613A	Forward
	MutA-C1613T-R		CACGGTGGT**T**TCCATGCAACG	1622	1602	C1613T	Reverse
	ACoreP-Mut-F		GGGGGAGGAGATTAGGTTAAT**G**A**T**CTTT	1742	1769	A1762T/ G1762A	Forward
	ACoreP-Mut-R		CAAAGA**T**C**A**TTAACCTAATCTCCTCCCC	1770	1743	T1762A/ C1764T	Reverse

The sites of mutation are indicated as red colour in bold.

Table 2.5 Primers used in site-directed mutagenesis in HBV core promoter.

2.4 Construction of 1.3-fold Full-length HBV Genomes of HBV Subgenotype Cs

The serum HBV DNA of subgenotype Cs from a chronic carrier was used as the template. As part of the work of a former M.Phil student, Miss Sophie Chan, our labpratory has successfully cloned a replicative competent 1.3-fold full-length HBV into pUC18 vector (designated as clone 'pG1'). A schematic diagram (Fig. 2.1) indicates the organization of the construct. The 1.3-fold HBV clone was designed to follow the organization of the 1.5-fold HBV clone (genotype A) from VIDRL (designated as clone 'PA'), only with a shorter genome length. It started from the X promoter and extended more than one genome length to include two copies of X ORF. It ended at 80 bp downstream of the poly A signal to include the *cis*-regulatory elements for effective replication.

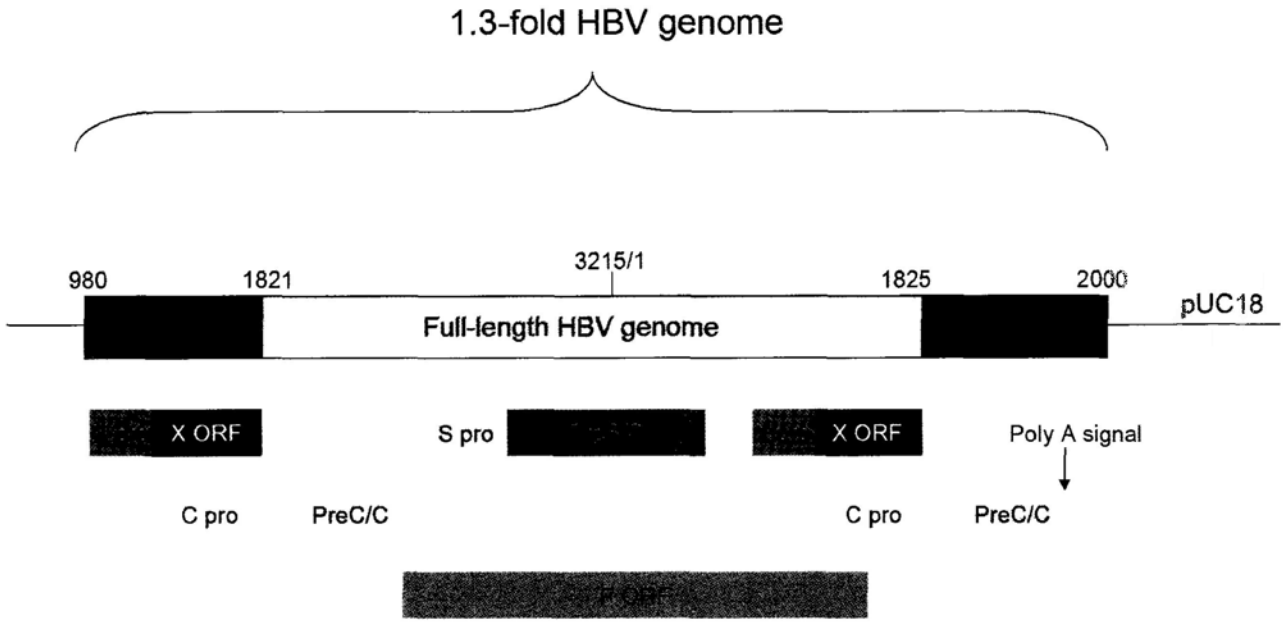


Fig. 2.1 The schematic diagram of the organization of 1.3-fold HBV construct. P: polymerase; pro: promoter; ORF: open reading frame; PreS/S: preS/S region; PreC/C: preC/C region.

2.4.1 Site-directed mutagenesis of 1.3-fold HBV genomes

Since the 1.3-fold genome spanned from nt. 980 to nt. 2000, which implied that the nt. 1613 appeared in the genome twice. This makes it more complicate to introduce mutations into the repeated region of the genome. Therefore, the genome needs to be separated into 2 parts for site-directed mutagenesis. Schematic diagrams are shown to illustrate the strategy of cloning as follows.

2.4.2 Introducing G1613A mutation in fragment nt. 980-2617

A 1,000-fold diluted 1.3x HBV genome construct (pG1) was used as template for the PCR as follows: 0.2 μM each of forward and reverse primers, 2 μl of diluted DNA template, 2× PicoMaxx® master mix (Stratagene) in a total volume of 25 μl. First, fragment nt. 980-1622 was generated using the forward primer (HindIII_980-F) and mutagenesis reverse primer (MutCC3Cor-R) and fragment nt. 1602-2617 was generated using mutagenesis forward primer (MutCC3Cor-F) and reverse primer (SacI-2617-R). Each of them was overlapped with a few bases including the target site of mutation (G1613A). The PCR condition was as follow: 94℃ for 36 s, 55℃ for

45 s and 72℃ for 2 min for 30 cycles and a final extension at 72℃ for 10 min. The PCR products were loaded to 1% TAE agarose gel and desired bands were purified by gel extraction kit (Qiagen) according to manufacturer's instruction. One microlitre of each purified fragment was linked together by a second round PCR without any primers for 10 cycles of amplification (94℃ for 36 s, 55℃ for 45 s and 72℃ for 2 min) and a final extension at 72℃ for 10 min. During the extension period, HindIII_980-F and SacI-2617-R were added and PCR resumed for 20 more cycles. This extended fragment contains an intrinsic *Apa*I site at nt. 2598. The final PCR products were loaded to 1% TAE agarose gel and desired bands were purified by gel extraction kit (Qiagen). After purification, the fragments were digested by *HindIII* and *ApaI* in Buffer M and Buffer L respectively at 37℃ for 16 hours. Then it was subsequently ligated into *Hind*III/*Apa*I digested pG1 plasmid for mini-scale purificaition (Qiagen) according to manufacturer's instruction. This clone, p1.3xHBV_980_2617_A, was used as a template for an additional mutation as described below. The schematic diagram of the mutagenesis is indicated in Fig. 2.2. The primers used are listed in Table 2.6.

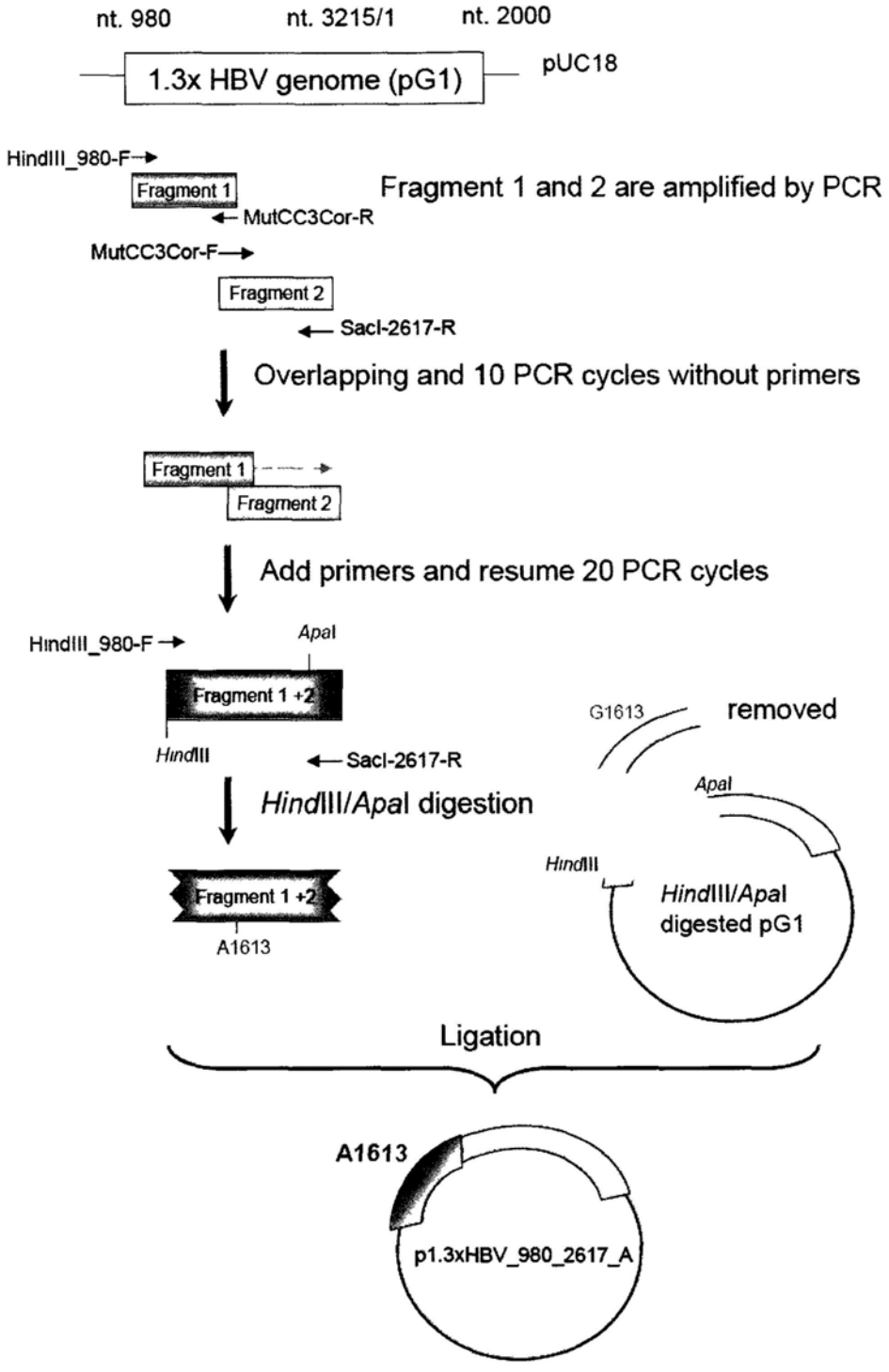


Fig. 2.2 Schematic diagram illustrating the cloning strategy to generate the first G1613A mutation in the 1.3-fold HBV genome.

2.4.3 Introducing G1613A mutation in fragment nt. 905-2000

Similarly, fragment nt. 905-1622 was generated using the forward primer (KpnI-905-F) and mutagenesis reverse primer (MutCC3Cor-R) and fragment nt. 1602-2000 was generates using mutagenesis forward primer (MutCC3Cor-F) and reverse primer (SacI-2000-R). They were then overlapped with a few bases including the target site of mutation (G1613A) and linked together by a second round of PCR without any primers for 10 cycles of amplification (94℃ for 36 s, 55℃ for 45 s and 72℃ for 2 min) and a final extension at 72℃ for 10 min. During the extension period, KpnI-905-F and SacI-2000-R were added and PCR resumed for 20 more cycles. The PCR products were loaded to 1% TAE agarose gel and desired bands were purified by gel extraction kit (Qiagen) according to manufacturer's instruction. After purification, the fragments were digested by *Kpn*I and *Sac*I in Buffer L at 37℃ for 16 hr, and then ligated into *Kpn*I/*Sac*I digested p1.3xHBV_980_2617_A plasmid for mini-scale purification (Qiagen). This clone was designated as pA1, which harbored G1613A mutations in the repeated core promoter regions. The schematic diagram of the mutagenesis is indicated in Fig. 2.3. The primers used are listed in Table 2.6.

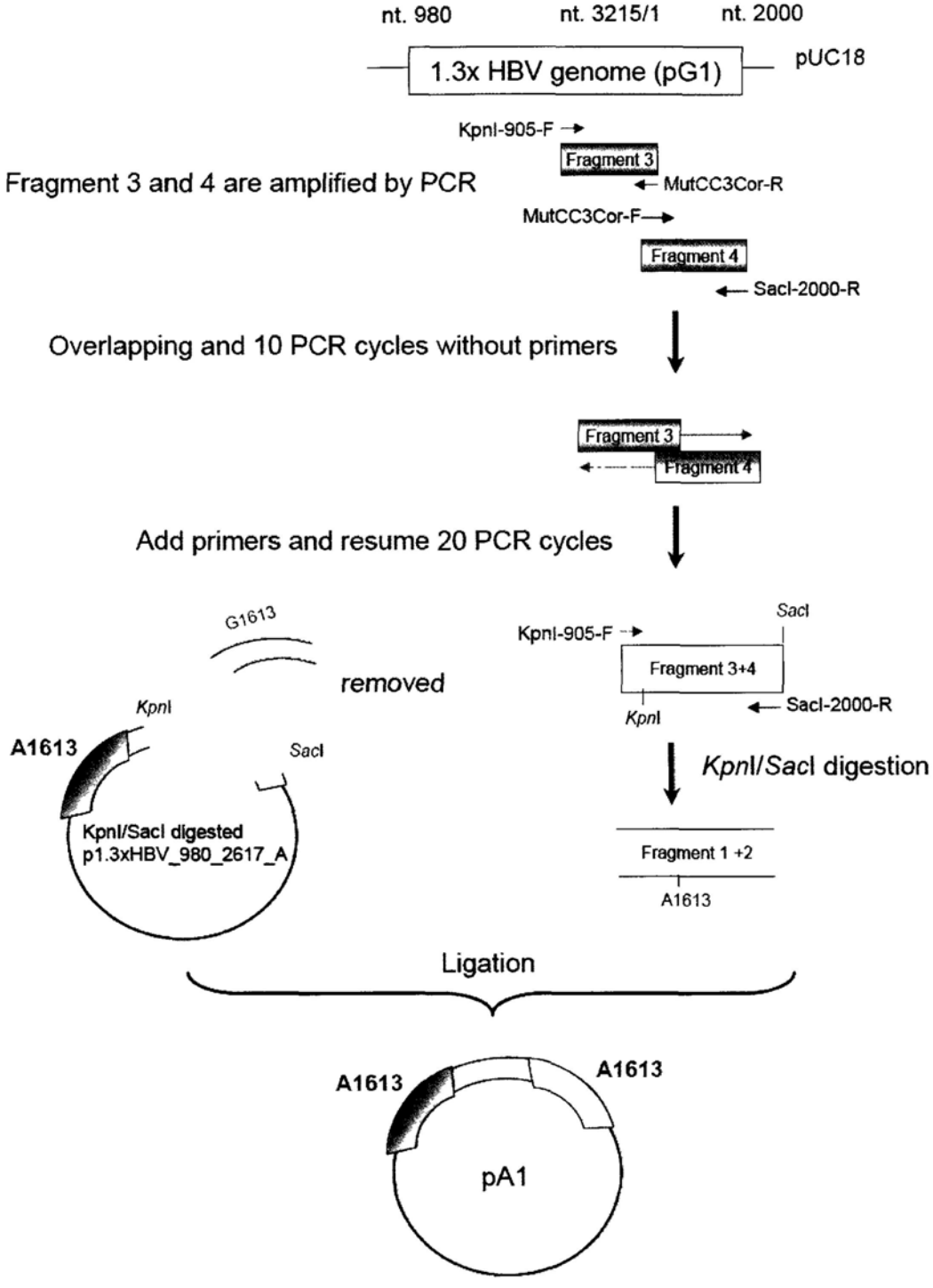


Fig. 2.3 Schematic diagram illustrating the cloning strategy to generate the second G1613A mutation in the 1.3-fold HBV genome.

Name	Sequence (5′ to 3′)	Start	End	Direction	Restriction site
SacI-Nt2617-R	GCGAGCTCTTTTCGTTAAC TGTAAGAGGGCCCACATA	2617	2589	Reverse	*Sac*I
KpnI-905-F	AAGGTTGGGGTACCTTGC CACAG	897	919	Forward	*Kpn*I
HindIII-980-F	GCAAGCTTGGAAGTATGT CAAAGAATTGTGGGTC	980	1005	Forward	*Hind*III
SacI-N2000-R	GCGAGCTCCGGTGTCGAG GAGATCACGAAT	2000	1978	Reverse	*Sac*I

Table 2.6 Additional primers used in PCR amplification of HBV fragments from 1.3x HBV construct.

2.4.4 Introducing the BCP back mutations in the p1.3x genomes

The G1613A mutation and the BCP mutation were introduced into the genome by site-directed mutagenesis in a strategy similar to that in described in section 2.4.2 and 2.4.3. The primers used in mutagenesis of the BCP back mutation are listed in Table 2.7. The sequences of all the full-length HBV clones and specific mutations were confirmed by DNA sequencing by using primers listed in Table 2.8.

Name	Sequence (5′ to 3′)	Start	End	Direction	Restriction site
C4-BCP-WT-F	TGGGGGAGGAGATTAGG TTAA**A**G**G**TCTTTG	1741	1770	Forward	*Sac*I
C4-BCP-WT-R	CAAAGA**C**C**T**TTAACCTAA TCTCCTCCCCCA	1770	1741	Reverse	*Kpn*I

Site of mutations are indicated as red colour in bold.

Table 2.7 Primers used in mutagenesis of BCP back mutation.

Name	Sequence (5′ to 3′)	Position	Direction
Seq2	TTGGCCAAAATTCGCAGTC	300-318	Sense
2356	CCCCACTGTTTGGCTTTCAG	715-734	Sense
970-991	GACATACTTTCCAATCAATAGC	970-991	Sense
3026	GCTGACGCAACCCCCACTGG	1186-1205	Sense
PC5	TCGCATGGAGACCACCGTGA	1607-1626	Sense
Core2ALT	CTCGGAACATTGTTCACCT	2031-2050	Sense
1075	AAGGTGGGAAACTTTACTGGGC	2469-2490	Sense
JM	TTGGGGTGGAGCCCTCAGGCT	3037-3057	Sense
1460-1441	GGGTCGTCCGCGGGATTCAG	1460-1441	Antisense
PC2	GGCAAAAACGAGAGTAACTC	1959-1940	Antisense
2076-2060	ATAGCTTGCCTGAGTGC	2076-2060	Antisense
903	GTTGATAAGATAGGGGCATTTGGTGG	2325-2300	Antisense
1798	CCACTGCATGGCCTGAGGATG	3219-3199	Antisense
M13U	GTAAAACGACGGCCAGT	Vector	Sense
M13R	CAGGAAACAGCTATGACC	Vector	Antisense

Table 2.8 Primers used in sequencing of 1.3-fold HBV genomes.

2.5 Cell Lines and Culture Medium

Three cell lines were used in this project. Human hepatoma cells HuH-7 from Japanese Collection of Research Bioresources Cellbank (JCBR) were cultured in Dulbecco's Modified Eagle Medium (DMEM) (Sigma-Aldrich) supplemented with 10% (v/v) heat-inactivated fetal bovine serum (FBS) (HyClone) and 1% streptomycin-penicillin (Invitrogen) at 37℃ in 5% CO_2 incubator. Culture of another two hepatoma cell lines, HepG2 and Hep3B cells, were similar but RPMI-1640 medium (Invitrogen Life Technology) was used.

2.6 Transient Transfection of Promoter Constructs into Hepatoma Cells

2.6.1 HuH7 cells

2.2 x 10^5 HuH-7 cells were seeded on 60-mm plate (Iwaki Glass) in 5 ml of DMEM to obtain approximately 60% confluence on the next day of transfection. The ratio of FuGENE6 (Roche Applied Science) to DNA used in HuH-7 cells had been

optimized as 8:5. Therefore, 4 μl of FuGENE6 was added directly to 100 μl of serum-free DMEM and the mixture was incubated for 5 min at room temperature. Then 2.5 μg of pGL3-promoter clone and 0.025 μg of pRL-CMV plasmid were added for transfection efficiency control. The complex was mixed and incubated for 20 min at room temperature and 100 μl of complex was added to each 60-mm dish in a drop-wise manner. Transfection of each type of pGL3-promoter clones were performed in triplicate. Cells were incubated at 37℃ in 5% CO_2 incubator for 2 days.

2.6.2 HepG2 and Hep3B cells

Transfection in 6-well plates was performed using the lipofectamine and PLUS reagents as per manufacturer's instruction (Invitrogen Life Technology). Lipofectamine and PLUS reagents were used instead of FuGENE6 because of the poor transfection efficiency of the FuGENE6 in these cell lines. In brief, 1 μg of pGL3-promoter construct and 0.01 μg of pRL-CMV were diluted in 100 μl plain medium and was pre-complexed with 6 μl of PLUS reagent at room temperature for 15 min. During the incubation, 4 μl of lipofectamine was diluted in 100μl of plain medium and the culture medium for the cells was replaced with plain medium. After incubation, the DNA mixture was combined with the diluted lipofectamine and

liposomes were allowed to form at room temperature for 15 min. The mixture was then added into the cells and the transfection was allowed to take place in a humidified CO_2 incubator for 5 hr. After that, cells were incubated with 3 ml new complete medium at 37℃ in 5% CO_2 incubator for 2 days.

2.6.3 Dual-luciferase reporter assay system

Cells were harvested at 2-day posttransfection. After washing with phosphate buffer saline (PBS), cells were lysed with 400 μl of 1X Passive Lysis Buffer (Promega) for 15 min at room temperature. Cell lysates were centrifuged at 13,000 rpm for 30 s to remove cell debris and nuclei. Luciferase assay was performed according to manufacturer's instruction. 5 μl of cell lysate was added with 100μl LARII (Promega) in 12 x 75 mm PYREX borosilicate glass tubes (Cornings Incorporation), which was then immediately placed into the single-tube luminometer, Lumat LB 9501, for the measurement of luminescent from Firefly luciferase at a 5-second measurement period. Afterwards, 100 μl of Stop & Glo Reagent (Promega) was promptly added to quench the Firefly signal and simultaneously activate the Renilla luciferase. A second luminescent signal was measurement for Renilla luciferase as the transfection efficiency control. Results were expressed as fold

increase relative to the corresponding control, which was set as 1. Data were reported as means ± standard deviations (S.D.) of 3 independent experiments, with triplicate in each experiment. The means between the control and mutant were compared by student *t* test. Analysis was performed using SPSS 11.5 or updated version. A *P* value < .05 was considered statistically significant.

2.7 Transient Transfection of 1.3-fold HBV Genomes in HuH7 Cells

Cells were seeded on 60-mm dishes in 5 ml of DMEM to obtain approximately 60% confluence on the next day for transfection. Eight microlitres of FuGENE 6 (Roche) was added directly to 100 μl serum-free OPTI-MEM (Invitrogen) and incubate for 5 min at room temperature. The diluted FuGENE6 was then added to 5 μg of 1.3-fold HBV genome, together with 0.1 μg of pSEAP-Control plasmid as transfection efficiency control. After incubation at room temperature for 20 min, the complexes were added to the cells and incubated at 37℃ and 5% CO_2. Growth media were changed 2 days after transfection. Cells and medium were harvested at 5-day post-transfection. The transfection efficiency was normalized by the secreted alkaline phosphatase (SEAP) activity in the culture media.

2.7.1 Extraction of HBV DNA from core particles

2.7.1.1 Intracellular core particles

At 5-day post-transfection, intracellular HBV DNA was prepared by adding 800 μl lysis buffer (10 mM Tris pH 7.5, 1 mM EDTA, 50 mM NaCl and 0.5% Nonidet P-40) at 4℃ for 15 min. Cells lysates were spun at 13,000 rpm for 5 min to remove the nuclei. Transfected plasmid DNA was eliminated from cell lysate by treatment with DNaseI (20U) and $MgCl_2$, and incubated at 37℃ for 2 hrs. Then followed by adding 5X stop buffer [2.5% (w/v) sodium dodecyl sulphate (SDS), 100 mM Tris pH 7.5 and 125 mM EDTA] inactivation and 15.6 mg/ml proteinase K (Invitrogen) digestion, which gave a final concentration of 0.5 mg/ml. The samples were incubated at 37℃ overnight for complete digestion of the core particles and release of HBV DNA. Finally, the intracellular HBV DNA was extracted using Qiamp DNA mini kit (Qiagen) according to manufacturer's instructions. DNA was eluted into 50 μl of nuclease-free water.

2.7.1.2 Extracellular HBV core particles

Extracellular HBV DNA was collected in cell culture media at 5-day post-transfection. The core particles in the media were precipitated by adding half volume of 26% polyethylene glycol buffer (Sigma) and incubated at 4℃ overnight. Then it was spun at 14,000 rpm for 20 min at 4℃. The pellet was then resuspended in 200 μl virus buffer (10 mM Tris pH 7.5 and 5.5 mM $MgCl_2$) and inactivate DNase I activity by adding 5X stop buffer. After that, the sample was subjected to proteinase K treatment as the same as that in intracellular HBV DNA preparation. Finally, the extracellular HBV DNA was extracted using Qiamp DNA mini kit (Qiagen) according to manufacturer's instructions. DNA was eluted into 50 μl of nuclease-free water.

2.7.2 Detection of HBsAg and HBeAg

Fifty microlitres of supernatant and cell lysate were collected for extracellular and intracellular HBeAg and HBsAg detection respectively by Microparticle Enzyme Immunoassay technology (AUTOBIO Diagnostics Co. Ltd). Mouse monoclonal anti-HBeAg antibody against HBeAg or anti-HBsAg antibody against HBsAg was

immobilized on the 96-well microtitre plate. After adding samples, the plate was incubated with enzyme conjugate at 37℃ for 30 min. After washing 5 times with wash buffer provided by the kit, 50 µl of substrate A followed by 50 µl of substrate B was added, and immediately incubated at 37℃ for 10 min in dark. Then the reaction was stopped by stop solution and the colour developed was measured at wavelength 450nm by a microplate reader. All readings exceeded the cut off value were interpreted as positive. They were normalized to the result of SEAP reporter assay as transfection efficiency control.

2.7.3 SEAP reporter gene assay

The SEAP Reporter Gene Assay (Roche Applied Science) was performed according to the manufacturer's instruction. Briefly, 50 µl of culture medium was diluted at a ratio of 1:4 with dilution buffer and incubated at 65℃ for 30 min. After centrifugation at 13,000 rpm for 30 s, 50 µl of sample was added to a 96-well opti-microplate in triplicates, followed by 5 min incubation with 50 µl inactivation buffer at room temperature. Finally, 50 µl of Substrate Reagent was added and incubated for

10 min at room temperature with gentle shocking. The light signal was detected by a multiplate reader.

2.7.4 Real-time quantitative PCR of HBV DNA

Two microlitres of 50-folds diluted DNA extracted were used in 10 μl real-time PCR reaction mixture. The reaction was carries out using a commercial 2X SYBR-green power master mix and 7500 Fast Real-Time PCR System (Applied Biosystems). Intracellular and extracellular HBV DNA was detected by primers according to a previous reported protocol (Maimuna *et al.*, 2006). The primer sequences are listed in Table 2.9. Each test run included positive and negative controls. The reaction conditions were modified to 95℃ for 10 mins followed by 40 cycles of 95℃ for 15 sec and 60℃ for 1 min. The level of HBV DNA was expressed in relative to the samples transfected with the constructs containing a 'G' at nt. 1613 with or without the BCP mutation.

Name	Sequence (5′ to 3′)	Start	End	Direction
HBV_S_F2	GTGTCTGCGGCGTTTTATCA	379	398	Forward
HBV_S_R2	GACAAACGGGCAACATACCTT	476	456	Reverse

Table 2.9 Primers used to detect HBV DNA using real-time quantitative PCR.

2.7.5 Primer Extension Assay

Primer extension assay for preC RNA and pg RNA was performed using a commercially available kit (Primer extension system-AMV reverse transcriptase, Promega). Huh7 cells were lysed at 5-day posttransfection using Trizol reagent (Invitrogen Life Science), and total cellular RNA was extracted with chloroform, and precipitated with isopropanol. RNA concentration was determined by spectrometer and confirmed by running an aliquot in a 1% TAE agarose gel. Ten microgram of RNA was annealed at 58 °C for 20 min with an antisense oligonucleotide 5′-GGAAAGAAGTCAGAAGGCAA -3′ (positions 1974-1955) that had been end labeled with γ-32P ATP using T4 polynucleotide kinase. The annealed oligonucleotide was extended by AMV reverse transcriptase at 42 °C for 30 min. The product was heated at 90 °C for 10 min and separated in 5% polyacrylamide gel containing 7M urea in 1× TBE buffer. The *Hinf*I digested ϕx-174 DNA was end labeled with γ-32PATP and run in parallel to serve as molecular size markers. The gel was dried and radioactive signals were detected by autoradiography or phosphoimager.

2.8 Data Analysis

The data in studying viral load in clinical samples were analyzed by the software SPSS11.5 for Windows. Continuous variables were expressed as mean with standard deviations (S.D.), and compared by *t* test or Mann-Whitney *U* test as appropriate. Categorical variables were compared by Chi-square test or Fisher's exact test as appropriate. All statistical tests will be 2 sided, and *P* values less than .05 will be considered as statistically significant.

2.9 Alignment of HBV NRE Sequences of HBV Genotype C

The HBV sequences were obtained from National Center for Biotechnology Information (GenBank) database (http://www.ncbi.nlm.nih.gov/). A total of 805 HBV sequences were aligned within the NREBP consensus binding site (nt. 1611-1619). The counts of each nucleotide were calculated.

2.10 Negative Regulatory Element Binding Protein (NREBP)

2.10.1 Cloning of NREBP binding domain (BP15)

The DNA sequencing encoding the DNA-binding domain of NREBP, BP15 (nt. 3941 to 5638), was cloned into the pCMV-myc vector (Clontech) using 1:50 fold diluted Hep3B cDNA as template. The PCR reaction contained 25 μl of 2× PicoMaxx master mix (Stratagene), 0.2 μM of forward and reverse primers (listed in Table 2.10) in 50 μl reaction. The PCR was performed with a 3 min initial denaturation at 94℃, followed by 10 cycles of amplification (94℃ for 36 s, 53℃ for 30 s and 72℃ for 1min 15s) and a final extension at 72℃ for 10 min. Afterwards, 2 μl of the PCR product was re-amplified for 30 cycles in a reaction volume of 50 μl with components same as above. The purified PCR products were cleaved by 0.5 μl each *EcoR*I and *Sal*I (Amersham Biosciences) in 1X buffer H, incubated at 37℃ for 16 hr followed by enzyme inactivation at 85℃ for 15 min. After purification, the digested product was ligated into 2 μl of pGL3-basic vector with T4 DNA ligase in 10X T4 DNA ligase buffer in a total volume of 20 μl and incubated at 4℃ overnight. The plasmid was transformed into *E.coli* (DH5α) and the purified clones were sequenced with

vector primers for further confirmation, and prepared in a large-scale by Hi-speed™ Plasmid Midi Kit (Qiagen) according to manufacturer's instruction.

Primer	Sequence (5'to3')	Start	End	Length	Tm	Enzyme
Forward	GAATTCgggtgctggcagagagcattctggag	3941	3964	24	58	*EcoR*I
Reverse	GTCGAC*tca*ttccctagacttggatctgtgctttg	5638	5613	26	56	*Sal*I

Table 2.10 Primers used to clone BP15. A stop codon was added to the reverse primer 3′ to the *Sal*I site (Italic). The restriction enzyme sites were underlined.

2.10.2 Transient overexpression of BP15 in HuH7 cells

Transfection in 6-well plates was performed using the FuGENE6 reagent as per manufacturer's instruction (Roche Applied Science). In brief, 3 μl of FuGENE6 reagent was diluted in 100μl of plain medium and incubated at room temperature for 5 min. Then 1 μg of plasmid was added and the mixture was incubated for 20 min at room temperature. Afterwards, 100 μl of mixture was added into the cells and incubated at 37℃ in 5% CO_2 incubator for 2 days.

2.10.3 Total protein extraction

Cells in 60-mm plate were washed with 1× PBS twice. Then 500 μl of lysis buffer (2% SDS, 10% glycerol and 0.0625 M Tris-HCl, pH 8.0) supplemented with cocktail protease inhibitor (Roche Applied Science) was added and incubated on ice for 15 min. The lysate was then centrifuged at 12,000 rpm for 15 min at 4℃ and the supernatant was kept in a microcentrifuge tube at -80℃ for storage.

2.10.4 Protein quantification using bicinchonic acid solution

Three microlitres of protein lysate or BSA protein standard were added in duplicates into a 96-well plate and 200 µl of assay reagent (49 parts 100% BCA : 1 part 4% $CuSO_4.5H_2O$) was mixed with the protein samples, and the mixture was incubated at 37℃ for 30 min before the spectrophotometric measurement at 540 nm using the water as blank.

2.10.5 Western blot analysis of BP15

Total protein was extracted according to section 2.6.5. The overexpression of the N-terminal c-myc tagged BP15 protein was examined by Western blot analysis. Ten micrograms of protein was separated in 10% SDS-polyacrylamide gel and electroblotted onto a PVDF membrane (Millipore). The blot was incubated in 5% non-fat milk in TBST at room temperature for 30 min, and then primary rabbit anti-myc antibody (Santacruz, SC-789) at a dilution 1:1000 was added for incubation overnight at 4℃. The blot was washed in TBST for 40 min and incubated with secondary antibody goat anti-rabbit IgG conjugated to HRP at a dilution 1:5000 for 1 hr at room temperature. After that, the blot was washed in TBST for 40 min and the

signals were detected using the Enhanced Chemiluminescence Western Blot Kit (Amersham Bioscience). After detection, the antibody on the blot was washed away by stripping buffer (Bio-Rad) at 37℃ for 20 min with rocking, and then washed with TBST. Mouse anti-β actin IgG conjugated to HRP was added at a dilution 1:10000 in TBST and the blot was incubated at room temperature for 1 hr. Afterwards, the blot was washed in TBST for 40 min and signal were detected using the Enhanced Chemiluminescence Western Blot Kit.

2.10.6 Co-transfection of BP15 with HBV promoters

HuH7 cells were seeded on 24-well plate at a confluency of 60%. Then, 0.6 μl of FuGENE6 was diluted in a dropwise manner in 20 μl of plain medium and incubated at room temperature for 5 min. Afterwards, 0.1 μg of each of pCMV-BP15 plasmid and pGL3-HBV promoter plasmid and 0.001 μg of pCMV-PRL plasmid were added to the diluted FuGENE6 and incubated at room temperature for 20 min. Then, 20μl of mixture was added into cells and incubated at 37℃ in 5% CO_2 incubator for 2 days. Dual luciferase reporter assay was performed according to manufacturer's instruction (refer to section 2.5.3).

2.10.7 Real-time PCR of endogenous NREBP

Eighteen paired HCC and non-tumourous tissues cDNA were generously given by Professor Winnie Yeo, Department of Clinical Oncology of the Chinese University of Hong Kong. The relative quantification of gene expression between HCC and non-tumourous tissues was examined using Applied Biosystems 7500 Fast Real-time PCR system. The real-time PCR contained 5 μl of Power SYBR Green PCR Master Mix (Applied Biosystems), 1 μl of diluted cDNA template, 0.125 μM of forward and reverse primers (listed in Table 2.11). The PCR was performed at initial heating at 94℃ for 3 min, followed by 40 cycles of 94℃ for 15 s and 60℃ for 30 s. The gene expression levels were normalized to an endogenous control gene, glyceraldehyde-3-phosphate dehydrogenase (GAPDH).

Gene Target	Primer	Sequence (5'to3')
BP15	Forward	TGGAACCAGCGGTGTCAGT
	Reverse	TGGACTCTATAGCCACCTCAGTTG
GAPDH	Forward	CGCCCCACTTGATTTTGA
	Reverse	TTGCCATCAATGACCCCTTCA

Table 2.11 Primers used in real-time quantitative PCR of BP15 and GAPDH.

2.11 Regulatory Factor X1 (RFX1)

The cDNA clone of Homo sapiens regulatory factor X1 (RFX1) was purchased from OriGene Technologies, Inc. (catalogue no. RC207872). The plasmid has been derived from single clone *E.coli* cultures and purified through ion-exchange chromatography. The eluant was normalized via UV spectroscopy, aliquoted into a Matrix vial as 10 μg transfection-ready dried plasmid DNA. The plasmid contained the full-length RFX1 ORF with a myc-DDK tag at the C-terminal cloned into the pCMV6 Entry vector (pCMV-RFX1). The sequence of this clone matched the reference sequence published in the National Center for Biotechnology Information with accession no. NM_002918.3 The DNA sequences of the RFX1 construct was confirmed by sequencing using primers listed in Table 2.12.

Primer	Direction	Sequence (5'to3')	Position	Priming site
RFX1-seqF1	Forward	CAGCCAGGTGCAGTATGTG	876-894	RFX1
RFX1-seqF2	Forward	GGAACACTGTGAGGCCATTG	1839-1858	RFX1
VP1.5	Forward	GGACTTTCCAAAATGTCG	-185	pCMV6 vector
XL39	Reverse	ATTAGGACAAGGCTGGTGGG	+199	pCMV6 vector

Table 2.12 Primer used in DNA sequencing of pCMV-RFX1 constructs.

2.11.1 Subcloning of pCMV-TNT-RFX1

The full-length ORF of RFX1 in pCMV-RFX1 plasmid was subcloned into pCMVTNT™ Vector (Promega), which is designed for the convenient expression of cloned genes using in vivo or in vitro expression systems. Both the SP6 and the T7 polymerase promoters lie in tandem adjacent to the multiple cloning sites, allowing for highly efficient synthesis of RNA in vitro from either promoter. Protein can be expressed in vitro from a gene cloned into the pCMVTNT™ Vector using an SP6- or T7-based, coupled in vitro transcription/translation system. The pCMVTNT™ Vector contains a 5′ β-globin leader sequence reported to enhance expression of certain genes in vitro. For in vivo expression, the vector contains a CMV enhancer/promoter region that can allow strong constitutive expression in many cell types. A β-globin/IgG chimeric intron and a late SV40 polyadenylation site are located downstream of the enhancer/promoter region. The primers used in subcloning are listed in Table 2.13. The PCR reaction included 0.2 μM each of forward and reverse primers, 1 μl of 1000-fold diluted pCMV-RFX1 and 25 μl of 2× PicoMaxx mastermix in a total volume of 50 μl. The PCR was performed with a 3 min initial denaturation at 94℃, followed by 32 cycles of amplification (94℃ for 36 s, 58℃ for 30 s and

72℃ for 3 min) and a final extension at 72℃ for 10 min. The PCR products were examined in 1% TAE agarose gel and purified by gel extraction kit (Qiagen), followed by *Xho*I and *Mlu*I digestion in 1X NEB buffer 3 at 37℃ for 16 hr. After that, the digested fragment was ligated to *Xho*I/*Mlu*I digested pCMV-TNT empty vector and subsequently transformed to E.coli and purified. The DNA sequences of the RFX1 construct (pCMV-TNT-RFX1) was confirmed by sequencing using primers listed in Table 2.12 and 2.13.

Primer	Direction	Sequence (5'to3')	Restriction site
XhoI-RFXI-F	Forward	GCCTCGAG*ATG*GCAACACAGGCGTATACTGAG	*Xho*I
MluI-RFXI-R	Reverse	GCACGCGT*TTA*GCTGGAGGGCAGCGCC	*Mlu*I

Table 2.13 Primers used in subcloning of RFX1 to pCMV-TNT vector. A start or stop codon (Italic) was added to the reverse primer 3′ to the restriction sites, which are underlined.

2.11.2 *In vitro* transcription/translation of RFX1 protein

The RFX1 gene was *in vitro* transcribed and translated using TNT® Quick Coupled Transcription/Translation Systems (Promega) according to manufacturer's instruction. In brief, the reaction contained 40 µl of Quick TNT master mix, 2 µl of radioactive ^{35}S-methionine (1000 Ci/mmol) (Perkin Elmer) or 20 µM of methionine, and 1-5 µg of DNA in 50 µl of reaction. The mixture was incubated at 30℃ for 90 min and the labeled translation products were analyzed by 10% SDS-PAGE. Then, the gel was blotted on 3mm Whatman paper and dried at 75°C for 1 hr, followed by signal detection by autoradiography or phosphorimager.

2.11.3 Co-transfection of pCMV-RFX1 with HBV promoters

Refer to section 2.10.6. The plasmid, pCMV-RFX1, was used instead of the pCMV-BP15 in the experiment.

2.12 Electromobility Shift Assay (EMSA)

Nuclear and cytoplasmic extracts of HuH7 cells was prepared by NE-PER® Nuclear and Cytoplasmic Extraction Reagents (PIERCE Biotechnology) according to manufacturer's instruction. The protein was aliquoted and kept at -80℃ until use. Annealed G1613 (wild-type) and A1613 (mutant) NRE probes and a non-specific oligo were end-radiolabeled with ^{32}P by T4 kinase (GE Healthcare). Unincorporated nucleotide was removed by G-25 column (GE Healthcare) according to manufacturer's instruction. The DNA binding reactions contained 5-20 μg of protein, 1 μl of labeled probe, 2 μl of 10×binding buffer (500 mM NaCl, 100 mM TrisHCl, pH 7.5, 25 mM MgCl2, 40% Glycerol, 5 mM DTT), 1 μl of 1% NP40 and 1 μl of poly (dI · dC) (GE Medical Systems) in a reaction volume of 20 μl. The reaction was incubated at room temperature for 1 hr and then analyzed in 5% native polyacrylamide gel. After that, the gel was blotted to a filter paper and dried at 75°C for 1 hr, followed by signal detection by autoradiography or phosphorimager. To validate the specificity of the binding, up to 250 times of cold probes were added to the binding reaction before the addition of the labeled probes. The difference in the binding affinities was reflected by the relative intensity of the specific gel-shifted bands. The sequences of the NRE probes and non-specific oligo are listed in Table 2.14.

Probe	Name	Direction	Sequence (5'to3')
G1613 (wild-type)	NRE30-G-F	Sense	GCACGTCGCATGGA**G**ACCAC CGTGAACGCC
	NRE30-G-R	Antisense	GGCGTTCACGGTGGT**C**TCCATGCGACGTGC
A1613 (mutant)	NRE30-A-F	Sense	GCACGTCGCATGGA**A**ACCAC CGTGAACGCC
	NRE30-A-R	Antisense	GGCGTTCACGGTGGT**T**TCCATGCGACGTGC
Non-specific	NS30-F	Sense	GGAATTACGTGGCCACTCGAGGGAATTACG
	NS30-R	Antisense	CGTAATTCCCTCGAGTGGCCACGTAATTCC

The position of nt.1613 is indicated as red colour in bold.

Table 2.14 The sequences of the wild-type and mutant NRE probes and non-specific oligos used in EMSA.

Chapter 3 RESULTS

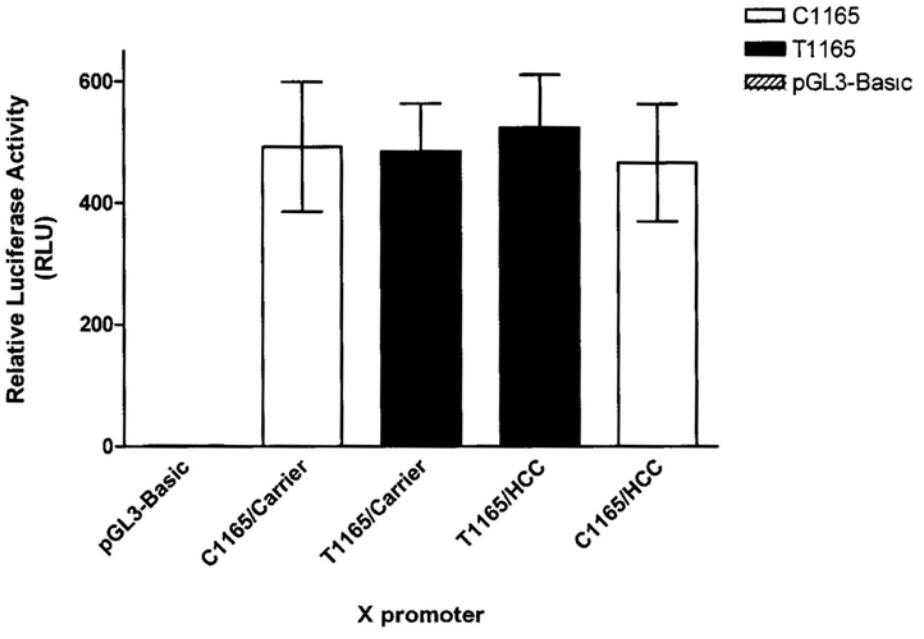
3.1 Effects of Three Promoter-related Mutations on HBV Genome

At the beginning of this study, we investigated the individual effect of 3 previously identified HCC-related mutations on HBV promoter regions. They are the C1165T mutation on X promoter, the T2712C mutation on preS1 promoter and the G1613A mutation on core promoter. For each mutation, 2 naturally occurring isolates each derived from the serum from a HBV carrier and a HBV-related HCC were obtained. They have the highest similarity to the HBV consensus sequence according to our database (see Chapter 2 Materials and Methods for details). For each isolate, a wild type and mutated promoter construct were cloned into pGL3-Basic vector (Promega) with a downstream Firefly luciferase gene. The corresponding promoter activities were examined by the luciferase reporter assay, and normalized to the co-transfected *Renilla* luciferase as the transfection efficiency control. Part of the work in the luciferase assay was finished by collaborating with Miss Sophie Chan, a former M.Phil student in my laboratory.

3.1.1 C1165T mutation does not alter HBV X promoter activity

Two HBV isolates of HBV genotype B derived each from carrier and HCC patient were obtained as templates. The configuration of the nt. 1165 in the isolate derived from the carrier is 'C' and we generated the mutant 'T' at this position by site directed mutagenesis. This pair of construct demonstrated the C1165T mutation on X promoter in the luciferase assay. Similarly, the configuration of the nt. 1165 in the isolate derived from the HCC patient is 'T', so it was mutated to 'C' to demonstrate the C1165T back mutation. The pGL3-basic was the empty vector control. The *Firefly* luciferase activities were normalized to the co-transfected *Renilla* luciferase as the transfection efficiency control. Our result showed that the C1165T mutation and the T1165C back mutation did not cause significant change in the X promoter activity in HuH7 cells (Fig. 3.1, upper panel). Similar result was obtained in HepG2 cells (Fig. 3.1, lower panel). The activity of the X promoter was strong with several hundreds folds greater than that of the promoterless pGL3-Basic vector. The positive control of the luciferase assay was the BCP mutation in core promoter (genotype A), which is indicated in Fig. 3.3.

The Effect of C1165T Mutation on HBV X Promoters Isolated from Carrier/HCC in HuH7 cells



The Effect of C1165T Mutation on HBV X Promoters Isolated from Carrier/HCC in HepG2 cells

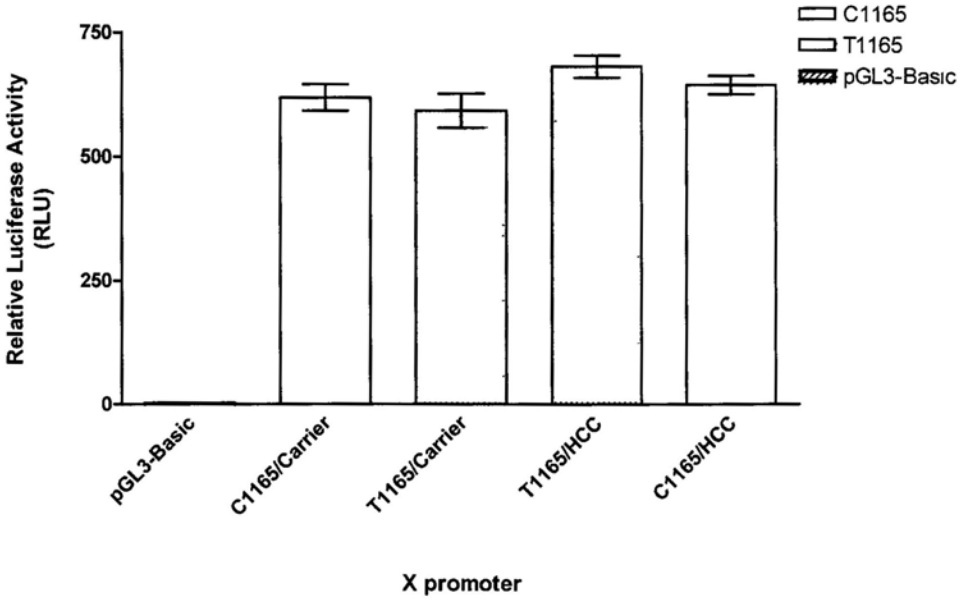
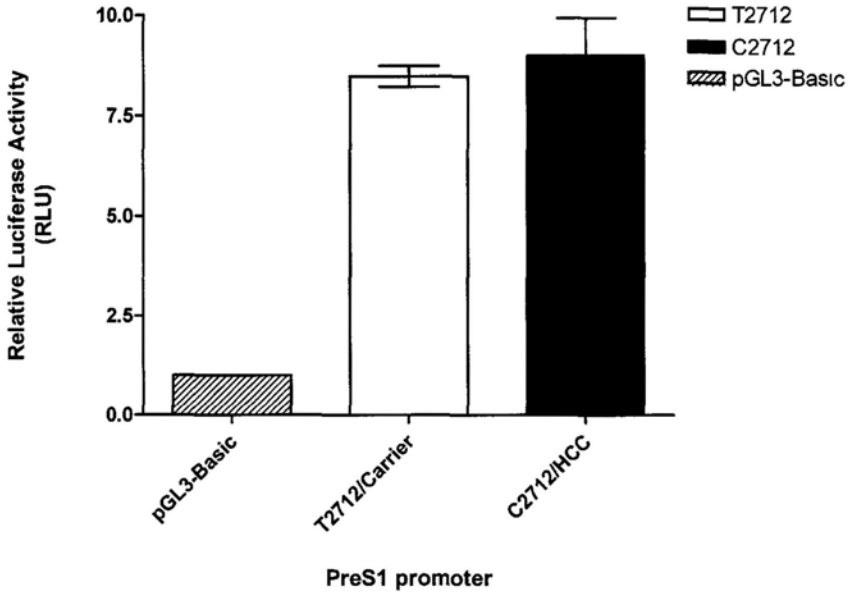


Fig. 3.1 The effect of C1165T mutation and T1165C back mutation on HBV X promoter. 1 μg of pGL3-promoter constructs and 0.01 μg pRL-CMV were transfected into hepatoma cells. Cells were harvested for luciferase assays at 2-day posttranfection. Results of luciferase assay in HuH7 (upper panel) and HepG2 cells (lower panel) are shown. The promoter constructs derived from the carrier demonstrated the C-to-T mutation, while those derived from the HCC patient demonstrated the T-to-C back mutation. The mutation did not alter the activity of the HBV X promoter. Results were presented as relative luciferase activity in 3 independent experiments with triplicates (mean ± S. D.) after normalizing to the co-transfected *Renilla* luciferase activities as transfection efficiency control.

3.1.2 T2712C mutation does not change HBV preS1 promoter activity

Since the sequence of the preS1 promoter from the isolates derived from the carrier and the HCC patient were the same except the site of mutation, only the promoter constructs obtained from the carrier's isolate and its mutant were examined. The isolate is belonged to HBV genotype B. The configuration of the nt. 2712 in the isolate derived from the carrier is 'T' and we generated the mutant 'C' at this position by site directed mutagenesis. This pair of construct demonstrated the T2712C mutation on X promoter in the luciferase assay. The pGL3-basic was the empty vector control. The *Firefly* luciferase activities were normalized to the co-transfected *Renilla* luciferase as the transfection efficiency control. As shown in Fig 3.2, the T2712C mutation did not affect the preS1 promoter activity in HuH7 and HepG2 cells. The activity of the preS1 promoter was relatively weak with only 10-20 folds greater than the promoterless pGL3-Basic vector. The positive control of the luciferase assay was the BCP mutation in core promoter (genotype A), which is indicated in Fig. 3.3.

The Effect of T2712C Mutation on HBV PreS1 Promoter Isolated from a HBV Carrier in HuH7 cells



The Effect of T2712C Mutation on HBV PreS1 Promoter Isolated from a HBV Carrier in HepG2 cells

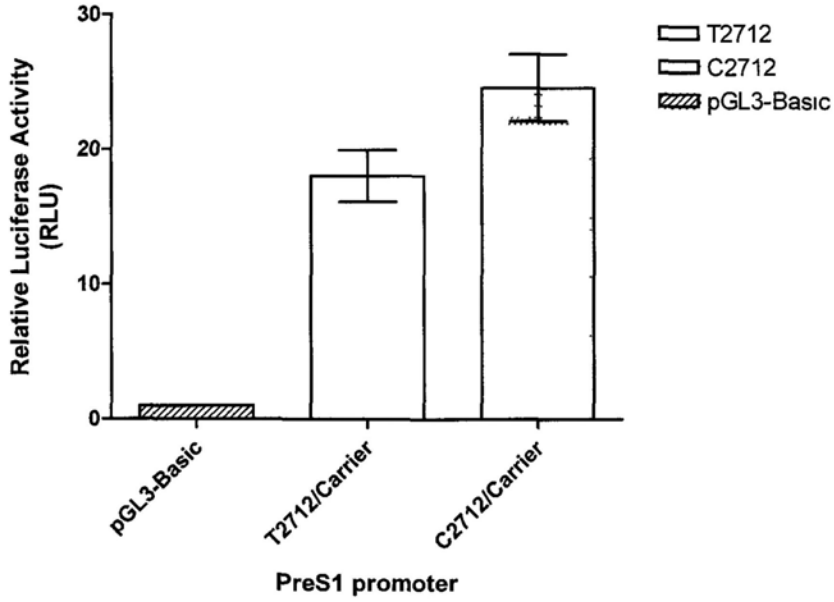


Fig. 3.2 The effect of T2712C mutation on HBV preS1 promoter. 1 μg of pGL3-promoter constructs and 0.01 μg pRL-CMV were transfected into hepatoma cells. Cells were harvested for luciferase assays at 2-day posttranfection. Results of luciferase assay in HuH7 (upper panel) and HepG2 cells (lower panel) are shown. Results of the luciferase assay in HuH7 (upper panel) and HepG2 cells (lower panel) are shown. The mutation did not show any significant effect on the activity of the preS1 promoter. pGL3-basic was the empty-vector control. Results are presented as relative luciferase activity in 3 independent experiments with triplicates (mean ± S. D.) after normalizing to the co-transfected *Renilla* luciferase activities as transfection efficiency control.

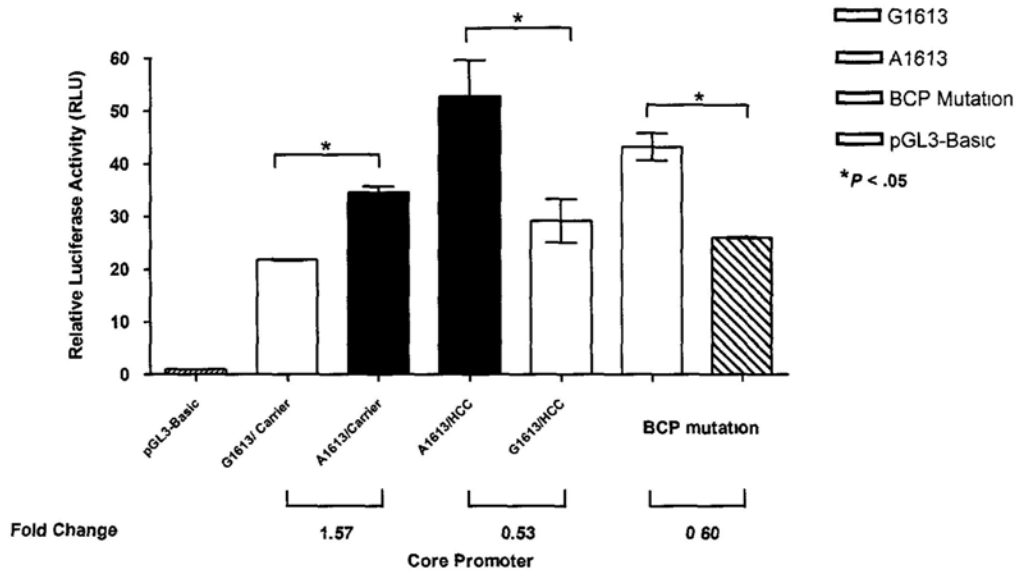
3.1.3 Effect of G1613A mutation on HBV core promoter activity

3.1.3.1 G1613A mutation enhances the core promoter activity

Two HBV isolates of HBV subgenotype Cs derived each from carrier and HCC patient were obtained as templates. The configuration of the nt. 1613 in the isolate derived from the carrier is 'G' and we generated the mutant 'A' at this position by site directed mutagenesis to demonstrated the G1613A mutation on core promoter in the luciferase assay. Similarly, the configuration of the nt. 1613 in the isolate derived from the HCC patient is 'A', so it was mutated to 'G' to demonstrate the A1613G back mutation. The pGL3-basic was the empty vector control. The *Firefly* luciferase activities were normalized to the co-transfected *Renilla* luciferase as the transfection efficiency control. In Figure 3.3 showed the G1613A mutation caused a significant increase in core promoter activity by 57% ($P < .003$) in HuH7 cells. The effect was reversed in the A-to-G back mutation, by 47% ($P < .003$). The constructs demonstrated the BCP mutation (genotype A) was included in the assay as a positive control, which caused 40% decrease in the core promoter activity ($P < .05$). Similar result was also obtained in HepG2 cells (Fig. 3.3, lower panel), in which the core promoter activity was significantly increased by 30% ($P < .02$). The effect was

reversed in the A-to-G back mutation by 30% ($P < .02$). Moreover, the BCP mutation decreased the core promoter activity by 22% ($P < .05$) which indicated the effect of the G1613A mutation is comparable to that of the positive control, despite their effects are opposite.

The Effect of G1613A Mutation on Core Promoters Isolated from Carrier/HCC in HuH7 cells



The Effect of G1613A Mutation on Core Promoters Isolated from Carrier/HCC in HepG2 cells

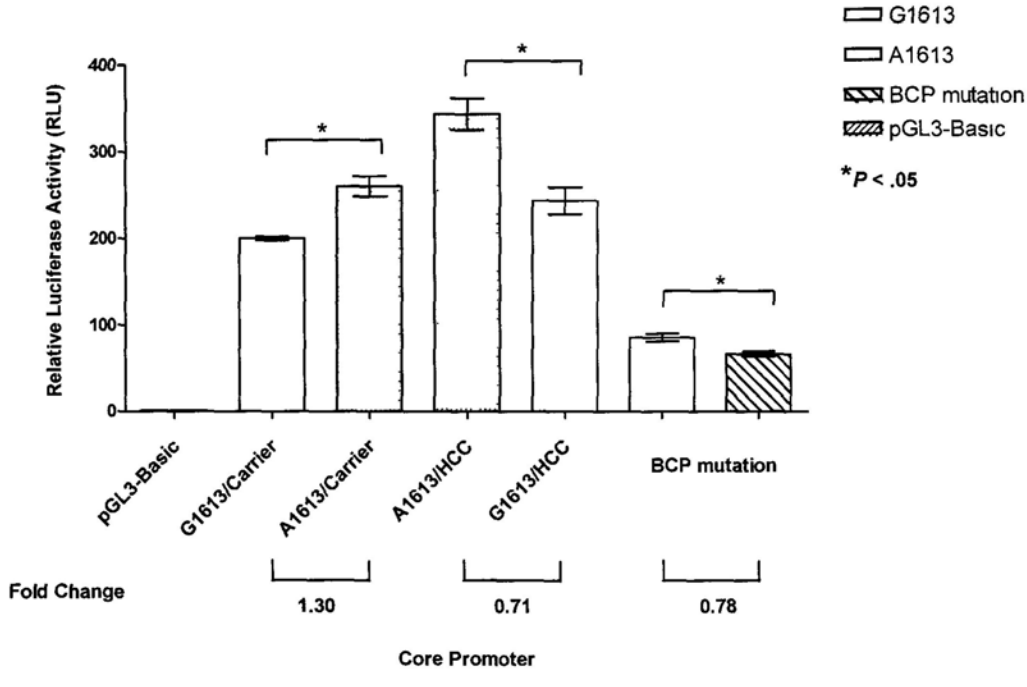
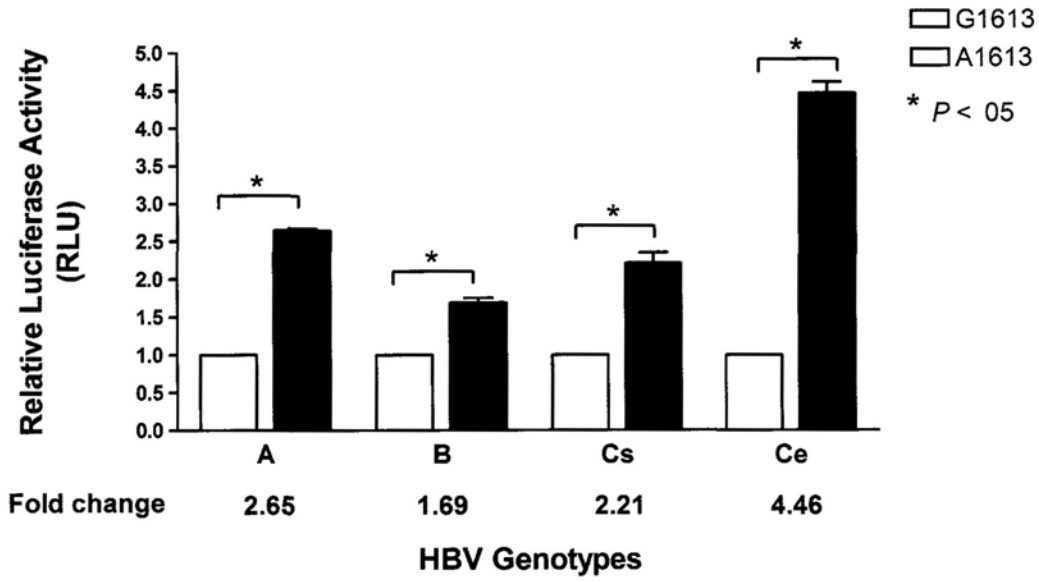


Fig. 3.3 The effect of G1613A mutation and A1613G back mutation on HBV core promoter. 1 μg of pGL3-promoter constructs and 0.01 μg pRL-CMV were transfected into hepatoma cells. Cells were harvested for luciferase assays at 2-day posttranfection. Results of luciferase assay in HuH7 (upper panel) and HepG2 cells (lower panel) are shown. The promoter constructs derived from the carrier demonstrated the G-to-A mutation, while those derived from the HCC patient demonstrated the A-to-G back mutation. The G1613A mutation significantly enhanced the promoter activities and *vice versa*. pGL3-basic was the empty-vector control. Results are presented as relative luciferase activity in 3 independent experiments with triplicates (mean ± S. D.) after normalizing to the co-transfected *Renilla* luciferase activities as transfection efficiency control.

3.1.3.2 G1613A mutation enhances core promoter activity in different HBV genotypes/subgenotypes

To investigate whether the G1613A mutation can enhance the core promoter activity in different HBV genotypes/subgenotypes, core promoters from different genotype background isolated from patients' serum were examined. Fig. 3.4 showed G1613A mutation increased the activities of core promoters in all genotypes and subgenotypes examined. The promoter activities were increased from 1.69 folds to 4.46 folds in HuH7 cells (Fig. 3.4, upper panel) ($P < .05$). Similar results were obtained in HepG2 cells (Fig. 3.4, lower panel). In contrast to the result in HuH7 cells, the G1613A mutation caused a weaker induction which ranged from 16% to 63 % in HepG2 cells ($P < .05$). Overall, the induction of promoter activity in genotype B was weaker than that in genotype C. Besides, the core promoter activity of subgenotype Ce was examined in Hep3B cells (Fig. 3.5), and it was induced by 2.8 folds by the G1613A mutation ($P < .01$).

Effect of G1613A Mutation on HBV Core Promoter in HuH7 cells



Effect of G1613A Mutation on HBV Core Promoter in HepG2 cells

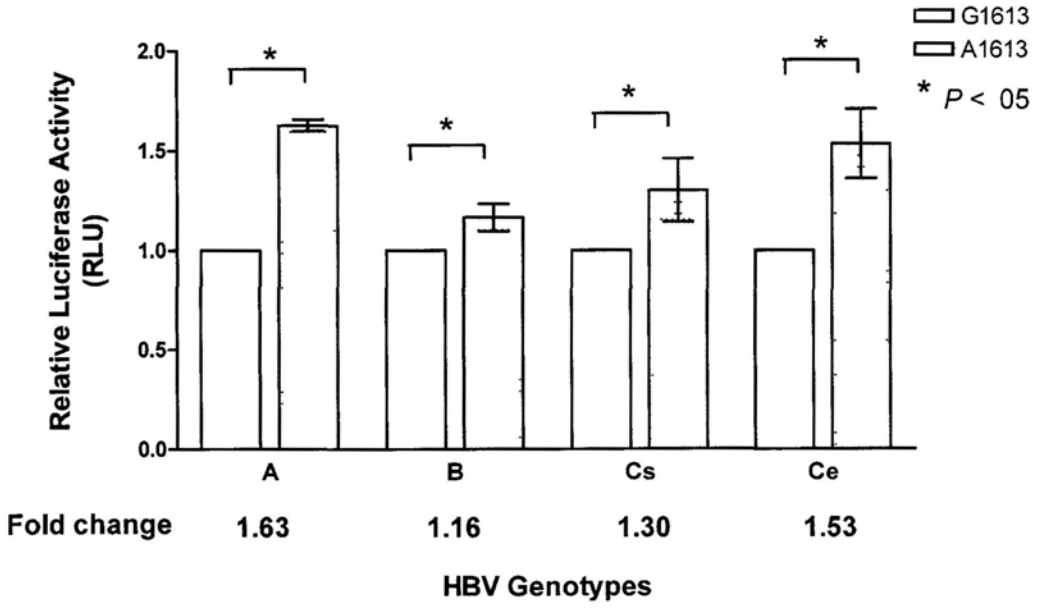


Fig 3.4 Effect of G1613A mutation on HBV core promoter of genotypes A and B, subgenotypes Cs and Ce. The HBV isolates were derived from patient's serum. They were all G1613 wild-type and mutated to A1613 by site-directed mutagenesis. 1 μg of pGL3-promoter constructs and 0.01 μg pRL-CMV were transfected into hepatoma cells. Cells were harvested for luciferase assays at 2-day posttranfection. Results of luciferase assay in HuH7 (upper panel) and HepG2 cells (lower panel) are shown. The mutation enhanced the core promoter activities in different HBV genetic background. Overall, the induction of promoter activity was weaker in genotype B than that in genotype C. Results were presented as relative luciferase activity in 4 independent experiments with triplicates (mean ± S. D.) after normalizing to the co-transfected *Renilla* luciferase activities as transfection efficiency control.

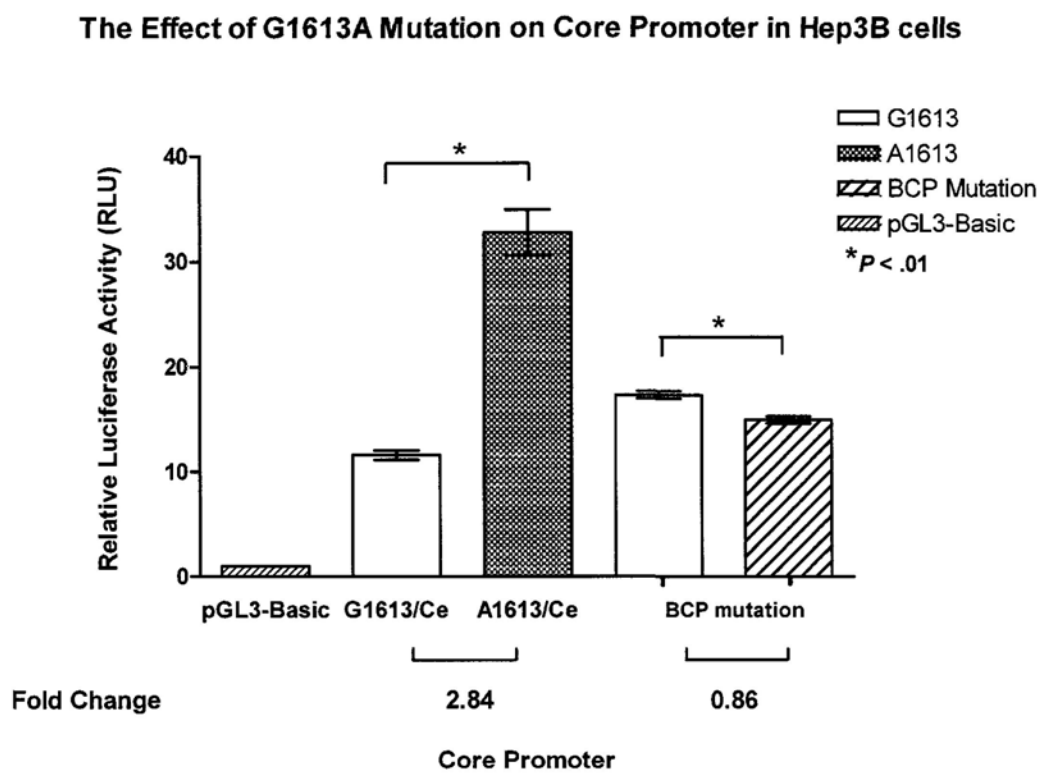


Fig. 3.5 The effect of G1613A mutation on core promoter activity of subgenotype Ce in Hep3B cells. The HBV isolate was derived from a patient's serum. The G1613A mutation was generated by site-directed mutagenesis. 1 μg of pGL3-promoter constructs and 0.01 μg pRL-CMV were transfected into HuH7 cells. Cells were harvested for luciferase assays at 2-day posttranfection. The mutation enhanced the core promoter activity by 2.84 folds. pGL3-Basic was the empty vector control and the BCP mutation was the positive control. Results are presented as relative luciferase activity in 4 independent experiments with triplicates (mean ± S. D.) after normalizing to the co-transfected *Renilla* luciferase activities as transfection efficiency control. Ce: HBV subgenotype Ce.

3.1.3.3 The effect of G1613A mutation was independent to that of the BCP mutation in subgenotype Cs

The BCP mutation is commonly found in HBV genotype C in both carriers and HBV-related HCC individuals (Orito *et al.*, 2001) Some studies suggested that the BCP mutation was associated to increased risk of HCC (Fang *et al.*, 2009; Kao *et al.*, 2003). In our study, the G1613A mutation increased the core promoter activity. Yet, the effects of the G1613A mutation and the BCP mutation on promoter activity are opposite. Therefore, it is worthwhile to test on the combined effect of the mutations on core promoter. In regard to this, two core promoter constructs, named pC1 and pC2, with a 'G' and 'A' at nt. 1613 respectively, and their BCP back mutated constructs, named pC3 and pC4 respectively, were included in our study. Figure 3.6 showed the BCP back mutation in subgenotype Cs caused a slight increase in promoter activity (16% in pC3/pC1 and 20% in pC4/pC2). In contrast, the G1613A mutation significantly increased the promoter activity (57% in pC2/pC1 and 64% in pC4/pC3). The effect caused by the G1613A mutation is about 3-fold greater than that of the BCP mutation. Interestingly, the activity in pC2 was significantly higher than that of pC3 (by 32%, $P < .03$). This suggests that the effect of G1613A mutation can overcome the effect of the BCP mutation.

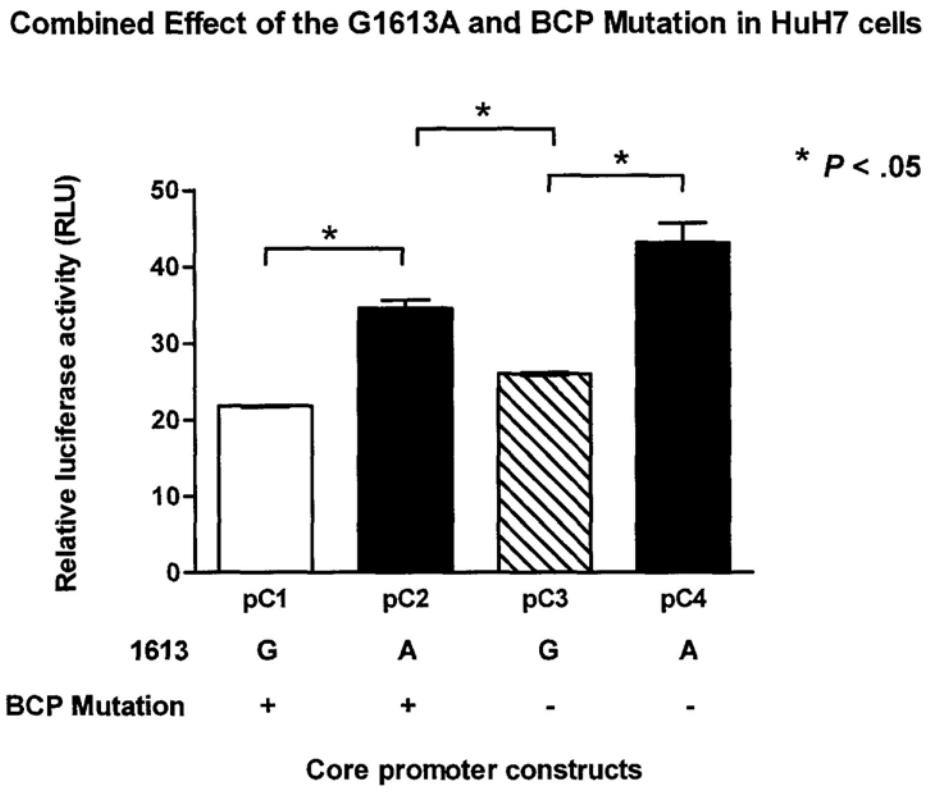


Fig. 3.6 The combined effect of the G1613A and BCP mutation on core promoter in HuH7 cells. The promoter constructs were derived from one patient's isolate of HBV subgenotype Cs. The constructs pC1 to pC4 contains the same core promoter sequence except the site of mutations as indicated. The G1613A and BCP mutation were generated by site-directed mutagenesis. 1 μg of pGL3-promoter constructs and 0.01 μg pRL-CMV were transfected into HuH7 cells. Cells were harvested for luciferase assays at 2-day posttranfection. The G1613A mutation increased while the BCP mutation decreased the promoter activity. The effect of the G1613A mutation was independent to that of the BCP mutation, with the enhancement caused by the G1613A mutation overcome the reduction caused by the BCP mutation (pC2 *vs* pC3). Results were presented as relative luciferase activity in 4 independent experiments with triplicates (mean ± S. D.) after normalizing to the co-transfected *Renilla* luciferase activities as transfection efficiency control.

3.2 G1613A is the Only Hotspot Mutation in the NRE of HBV Genotype C

The G1613A mutation is located in the NRE region, which has been shown to repress the activity of CURS region and enhancer II by 10-20 folds (Chen and Ou., 1995; Lo and Ting, 1994). In regard to this, we have obtained 803 HBV complete genome sequences reported in National Center for Biotechnology Information (Table 3.1). All the sequences are belonged to HBV genotype C. The NRE region including nt. 1604-1636 were selected for alignment (data provided in appendix). Our result showed that the nt. 1613 is the only hot spot mutation in the NRE within the nt. 1611-1619.. Among the 803 HBV sequences, 583 (72.6%) of them confers to 'G', 209 (26%) of them confers to 'A', while 11 (1.4%) of them confers to 'R' which defines purines.

Position	Nucleotide on HBV NRE region	Count	%
1611	G	803	100.0
1612	A	795	99.0
	C	8	1.0
1613	**G**	**583**	**72.6**
	A	**209**	**26.0**
	R	**11**	**1.4**
1614	A	798	99.4
	G	5	0.6
1615	C	803	100.0
1616	C	803	100.0
1617	A	799	99.5
	T	2	0.2
	C	2	0.2
1618	C	803	100.0
1619	C	803	100.0

Table 3.1 Analysis of 803 NRE sequences of HBV genotype C. The nucleotide 1613 is the only hot spot mutation within the NRE region (nt. 1611-1619). The HBV NRE sequences were obtained from National Center for Biotechnology Information (GenBank) database (http://www.ncbi.nlm.nih.gov/).

3.3 Prevalence of the G1613A Mutation in HBV Carriers and HBV-related HCC Individuals

In a statistical analysis of the prevalence of the G1613A mutation in 99 HBV carriers and 100 HCC patients, we found that the G1613A mutation was present in HBV isolates of genotype B and subgenotypes Cs and Ce (Table 3.2). The prevalence of the mutation increased from 22% (in carriers) to 35.1% (in HCC) in genotype B, and from 10% (in carriers) increased to 37.5% (in HCC) in subgenotype Ce. Despite of that, it was found as statistical insignificant. The reasons will be discussed in Session 4.2 in Chapter 4. In contrast, the mutation was significantly associated to HCC in subgenotype Cs. The prevalence of the mutation increased from 10.3% (in carriers) to 36.2% (in HCC) (P =.003).

HBV Genotype/ Subgenotype	Cases	Number of cases	Nucleotide at 1613 G(%)	A(%)	R (%)	*P* value
B	Carriers	50	39 (78.0)	11 (22.0)	0 (0.0)	.175
(N=87)	HCC	37	24 (64.9)	13 (35.1)	0 (0.0)	
Cs	Carriers	39	33 (84.6)	4 (10.3)	2 (5.1)	* .003
(N=86)	HCC	47	29 (61.7)	17 (36.2)	1 (2.1)	
Ce	Carriers	10	9 (90.0)	1 (10.0)	0 (0.0)	.179
(N=26)	HCC	16	9 (56.3)	6 (37.5)	1 (6.3)	

*The G1613A mutation was found significantly higher prevalence in HCC patients in subgenotype Cs.

Table 3.2 The prevalence of G1613A mutation in HBV carriers and HCC patients.

3.4 Prevalence of the G1613A Mutation to High Serum Viral Load in HBV Subgenotype Cs

In this study, we obtained the configuration of nucleotide 1613 on HBV core promoter from 255 chronic carriers. The quantification of HBV viral load and α-fetoprotein (AFP) level were determined. This part of work was assisted with Prof. Henry L. Y. Chan in Prince of Wales Hospital. The correlation of the HBV viral load and the prevalence of the G1613A mutation on the HBV genome were determined in a univariate analysis. As high serum viral load often indicates higher risk to HCC (Liu *et al.*, 2008), especially for vial load higher than 6 log copies/ml (Wong *et al.*, 2010). Therefore, the cut-off value of the viral load in our study was set as 6 log copies/ml. Among the HBV carriers, 129 patients (41 females and 88 males) had their serum viral load more than 6 log copies/ml. As shown in Table 3.3, the G1613A mutation was significantly associated to high serum viral load (more than or equal to 6 log copies/ml) in female carriers (OR = 3.657, P = .01). This indicated that the G1613A mutation may correlated to the high viral load in patients and increase the risk to HCC. Meanwhile, the relationship between the prevalence of G1613A mutation and serum AFP level was also investigated. However, we found no significant difference between the groups of wild-type G1613 and mutant A1613 (Table 3.4).

Cases (Total N=255)	Viral load (log copies/ml)	Nucleotide at 1613 G (%)	A (%)	*P* value	Harzard Ratio (95% CI)
Male (N =167)	< 6	62 (78.5)	17 (21.5)	.589	0.81 (0.378-1.737)
	≥ 6	72 (81.8)	16 (18.2)		
Female (N =88)	< 6	40 (85.1)	7 (14.9)	* .01	3.657 (1.32-10.133)
	≥ 6	25 (61.0)	16 (39.0)		

* The G1613A mutation is associated with higher (more than or equal to 10^6 copies/ml) serum viral load in female carriers.

Table 3.3 The relationship between the prevalence of G1613A mutation and viral load in HBV chronic carriers of subgenotype Cs

Cases (Total N=255)	Nucleotide at 1613	AFP Level (ng/ml) Mean (± S. E. M.)	*P* value (95% CI)
Male (N =167)	G (N=134)	3.74 (± 1.08)	.252 (-3.91-5.05)
	A (N=33)	3.15 (± 1.21)	
Female (N =88)	G (N=65)	6.04 (± 2.78)	.658 (-6.33-12.59)
	A (N=23)	2.91 (± 1.40)	

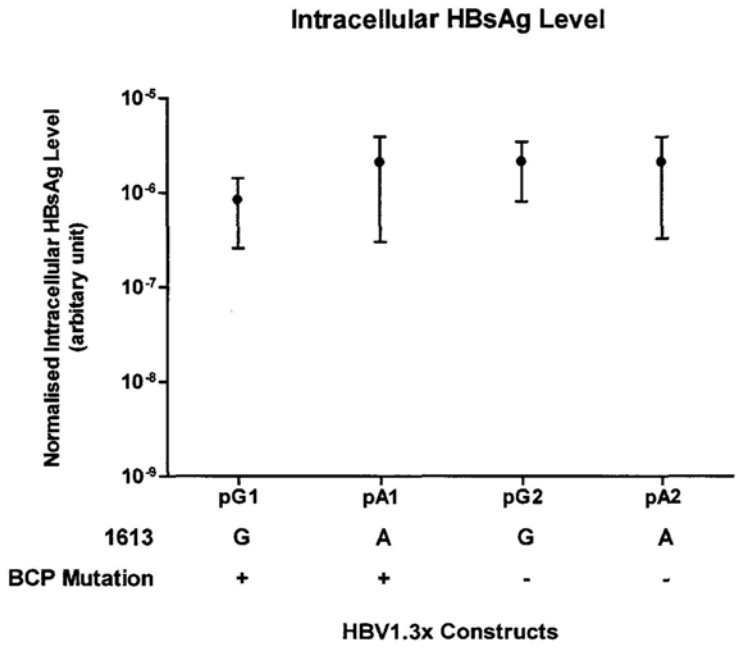
Table 3.4 The relationship between the prevalence of G1613A mutation and AFP level in HBV chronic carriers of subgenotype Cs.

3.5 *In vitro* Effect of the G1613A Mutation on Virus

3.5.1 G1613A mutation suppressed HBeAg production with no change on HBsAg level.

The effects of the G1613A mutation on viral HBsAg and HBeAg production were examined in a full-length HBV genome transfection assay. Four constructs of subgenotype Cs (pG1, pA1, pG2 and pA2) containing 1.3-fold HBV genome were used to test the single or combined effect of the G1613A and BCP mutation. We found that the intracellular and extracellular levels of HBsAg and the intracellular level of HBeAg were not affected by the mutations (Fig. 3.7 and 3.8A). In contrast, when compared with the corresponding 1613 wild-type constructs, the G1613A mutation significantly decreased the extracellular HBeAg level by 90% and 86%, in the presence and absence of the BCP mutation respectively (Fig. 3.8B).

(A)


Intracellular HBsAg Level
Normalised Intracellular HBsAg Level (arbitary unit)
10-5
10-6
10-7
10-8
10-9
pG1
pA1
pG2
pA2
1613
G
A
G
A
BCP Mutation
+
+
-
-
HBV1.3x Constructs

(B)

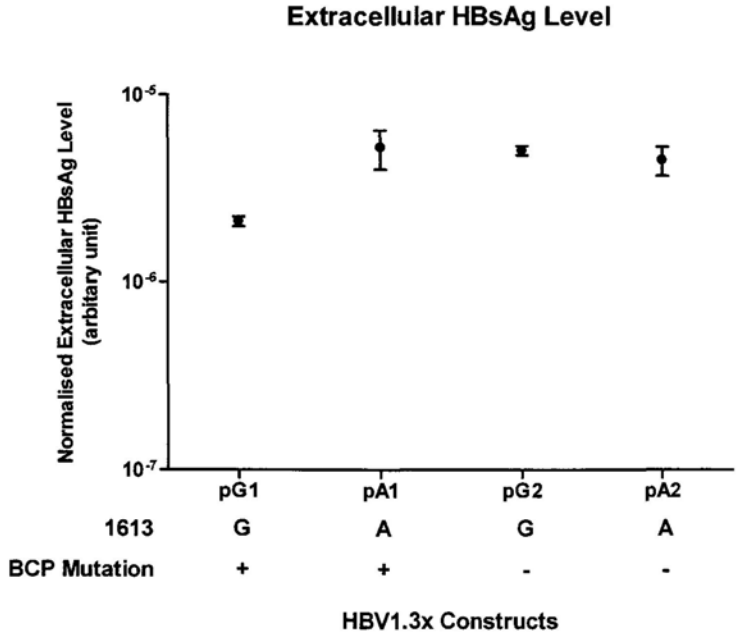
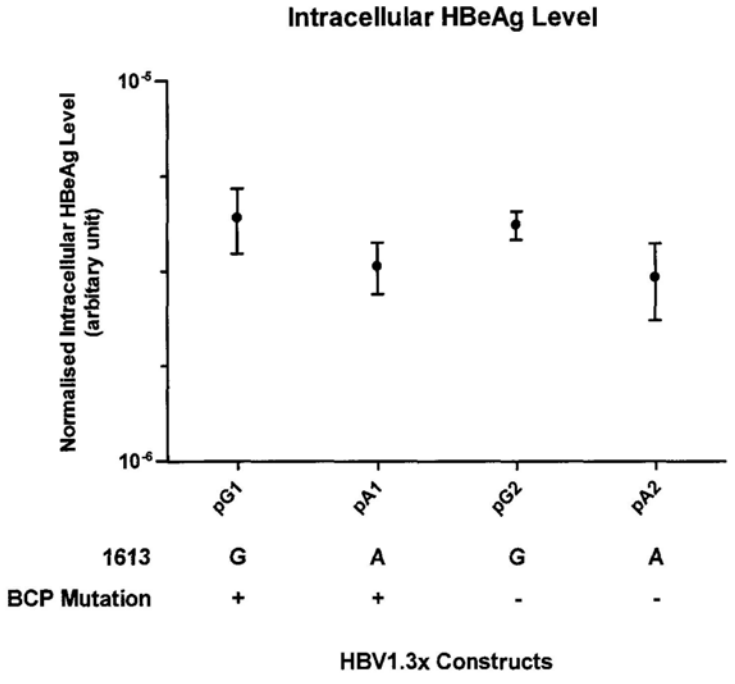

Extracellular HBsAg Level
Normalised Extracellular HBsAg Level (arbitary unit)
10-5
10-6
10-7
pG1
pA1
pG2
pA2
1613
G
A
G
A
BCP Mutation
+
+
-
-
HBV1.3x Constructs

Fig. 3.7 Intracellular (A) and extracellular (B) HBsAg level at 5-day posttransfection of 1.3x HBV plasmid into HuH7 cells respectively. All the constructs had the same genetic background except the mutations as indicated in the figures. The mutations were introduced into the HBV genome by site directed mutagenesis. 5 μg of 1.3-fold HBV genome was cotransfected with 0.1 μg of pSEAP-control plasmid. The relative amounts of intracellular and extracellular HBsAg were measured by ELISA. Result showed that the G1613A mutation did not cause any significant changes on the level of the HBsAg production when compared to the G1613 wild-type. Results were normalized to the activity of the co-transfected alkaline phosphatase (SEAP) activity as the transfection efficiency control. Data were presented as normalized values in 4 independent experiments (mean ± S. D.).

(A)


Intracellular HBeAg Level
Normalised Intracellular HBeAg Level (arbitary unit)
10^{-5}
10^{-6}
pG1
pA1
pG2
pA2
1613
G
A
G
A
BCP Mutation
+
+
-
-
HBV1.3x Constructs

(B)

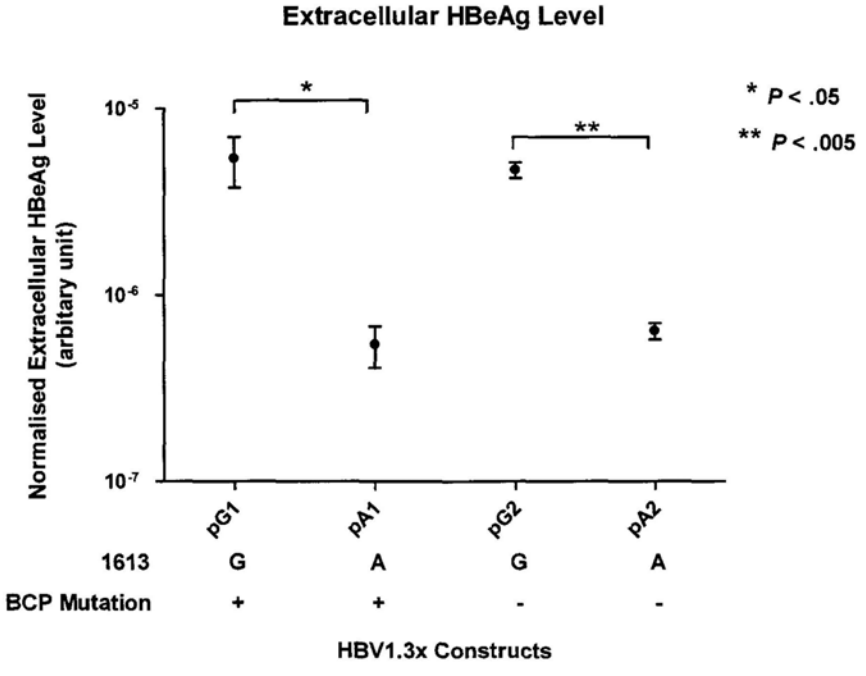
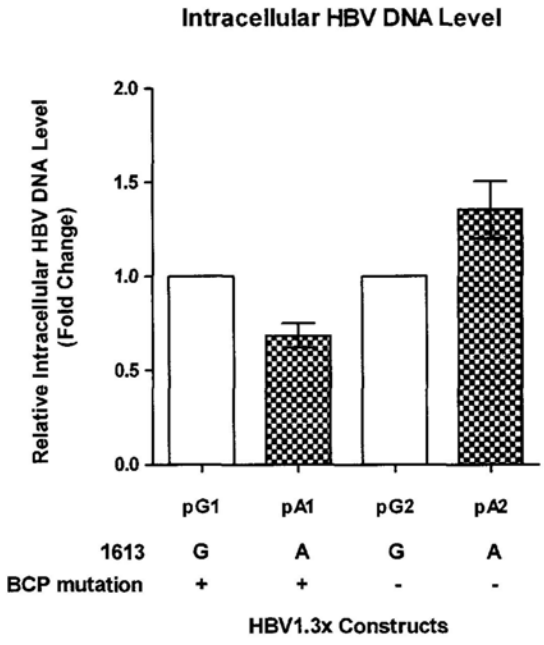

Extracellular HBeAg Level
Normalised Extracellular HBeAg Level (arbitary unit)
10^{-5}
10^{-6}
10^{-7}
*
**
* P < .05
** P < .005
pG1
pA1
pG2
pA2
1613
G
A
G
A
BCP Mutation
+
+
-
-
HBV1.3x Constructs

Fig. 3.8 Intracellular (A) and extracellular (B) HBeAg level at 5-day posttransfection of 1.3x HBV plasmid into HuH7 cells respectively. All the constructs had the same genetic background except the mutations as indicated in the figures. The mutations were introduced into the HBV genome by site directed mutagenesis. 5 μg of 1,3-fold HBV genome was cotransfected with 0.1 μg of pSEAP-control plasmid. The relative amounts of intracellular and extracellular HBeAg were measured by ELISA. Result showed that the G1613A mutation drastically dropped the level of the extracellular HBeAg, while no change on the intracellular HBeAg level when compared to the G1613 wild-type. Results were normalized to the activity of the co-transfected SEAP as the transfection efficiency control. Data were presented as normalized values in 4 independent experiments (mean ± S. D.).

3.5.2 Effect of the G1613A mutation on viral DNA production

To investigate the effect of G1613A mutation on HBV DNA production, 1613 wild-type and mutant 1.3-fold HBV genomes (pG1, pA1, pG2 and pA2) were transfected into HuH7 cells. The core-particle associated HBV DNA was collected and extracted in both cell lysate and culture media as the intracellular and extracellular HBV DNA respectively at 5-day posttransfection. The HBV DNA levels were determined by real-time quantitative PCR using primers specific to HBV S gene. The results were normalized to the HBV DNA level of the corresponding 1613 wild-type constructs, which were set as '1'. G1613A mutation did not alter the level of intracellular HBV DNA (Fig 3.9A). In contrast, Figure 3.9B showed the mutation significantly increased the level of extracellular HBV DNA by 2-fold and 4-fold in the presence and absence of the BCP mutation respectively ($P < .05$)..

A.


Intracellular HBV DNA Level
Relative Intracellular HBV DNA Level (Fold Change)
2.0
1.5
1.0
0.5
0.0
pG1
pA1
pG2
pA2
1613
G
A
G
A
BCP mutation
+
+
-
-
HBV1.3x Constructs

B.

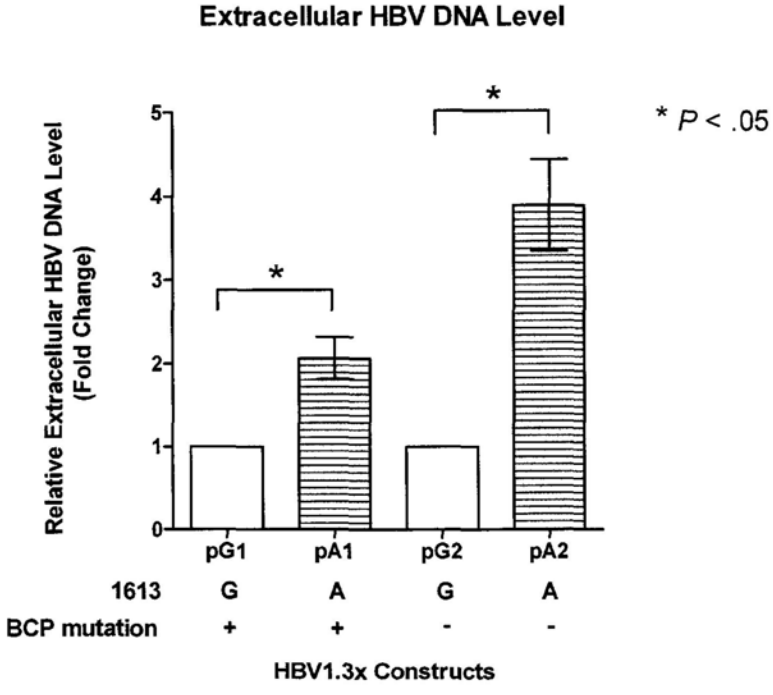

Extracellular HBV DNA Level
Relative Extracellular HBV DNA Level (Fold Change)
5
4
3
2
1
0
*
*
pG1
pA1
pG2
pA2
1613
G
A
G
A
BCP mutation
+
+
-
-
HBV1.3x Constructs
* P < .05

Fig. 3.9 The relative intracellular (A) and extracellular (B) HBV DNA level at 5-day posttransfection respectively. HuH7 cells were transfected with 5μg 1.3-fold HBV genomes. All the constructs had the same genetic background except the G1613A and BCP mutations as indicated in the figures. The G1613A mutation significantly increased viral DNA production to the extracellular media, while no significant change in the intracellular viral DNA level. Results were shown in mean values ± S.D. from 3 independent experiments, compared to the G1613 wild-type constructs which were set as '1' for comparison.

3.5.3 Effect of G1613A mutation on differential transcription of preC mRNA and pgRNA from core promoter

The HBV core promoter plays a very critical role on the HBV replication as it directs the transcription of the preC mRNA and pgRNA. The preC mRNA is translated to the precore protein, which is the precursor protein of HBeAg. The pgRNA is the template of the viral DNA synthesis. The preC mRNA and pgRNA are both 3.5kb long and they only differ by their 5′ end, which the preC mRNA is 30 nucleotides longer. To further investigate the functional effect on the G1613A mutation on virus RNA transcription, primer extension assay was performed to measure the relative amount of the two transcripts after the transfection of the 1.3-fold HBV genomes. Total RNA was extracted after transient transfection of 1.3-fold HBV genomes into HuH7 cells at 5-day posttransfection. As shown in Fig. 3.10, the integrity of the RNA remains after extraction and appeared as sharp bands of 18S and 28S ribosomal RNA in 1% TAE agarose gel stained with ethidium bromide.

The result of the primer extension assay is indicated in Figure 3.11A. The product of preC mRNA (145 bp) and pgRNA (115 bp) were successfully amplified in the experiment. Their relative signals were calculated by densitometer and

normalized to the SEAP reporter activity as the transfection efficiency control (Fig. 3.11B). As shown in Figure 3.11C, the preC mRNA level was significantly down-regulated by 53% (pA1 *vs* pG1) and 46% (pA2 *vs* pG2) by the G1613A mutation in the presence and absence of the BCP mutation respectively. Since precore protein was encoded from the preC mRNA, and it is the precursor protein for the HBeAg synthesis. Our results revealed that the G1613A mutation suppressed HBeAg production on the transcription level. In contrast, the BCP mutation did not caused significant change on preC mRNA level. This was consistence to the insignificant change on the HBeAg level (Fig. 3.8B). On the other hand, the pgRNA is the template for HBV DNA synthesis. Our result showed that the levels of pgRNA in the A1613 mutants were decreased by 21% (pA1 *vs* pG1) and 39% (pA2 *vs* pG2) in the presence and absence of the BCP mutation respectively (Fig. 3.11D). Interpretation of this observation would be discussed in details in Section 4.4.

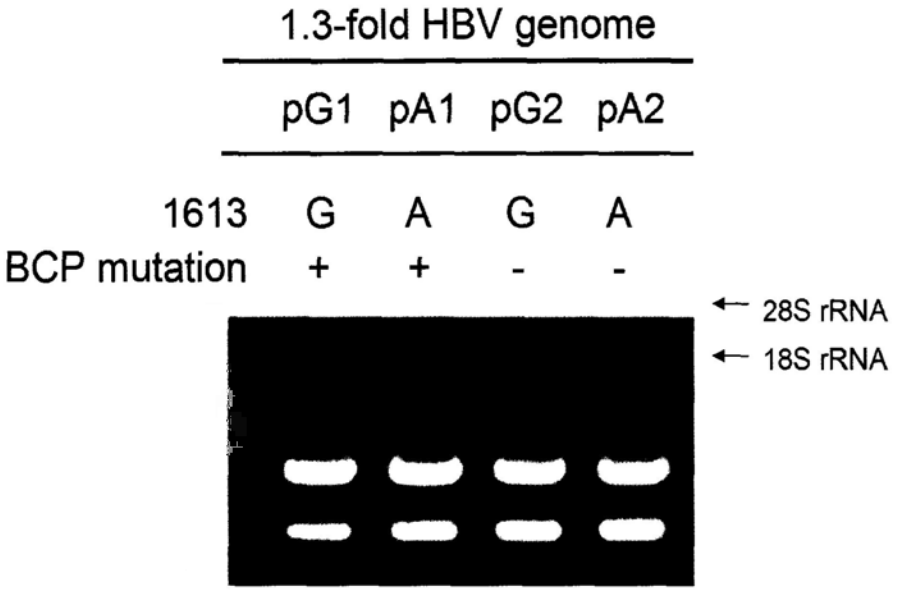


Fig. 3.10 Gel electrophoresis of total RNA extracted at 5-day posttransfection of 1.3-fold HBV genomes in HuH7 cells. Distinct bands of 28S and 18S rRNA showed the integrity of RNA on 1% TAE agarose gel.

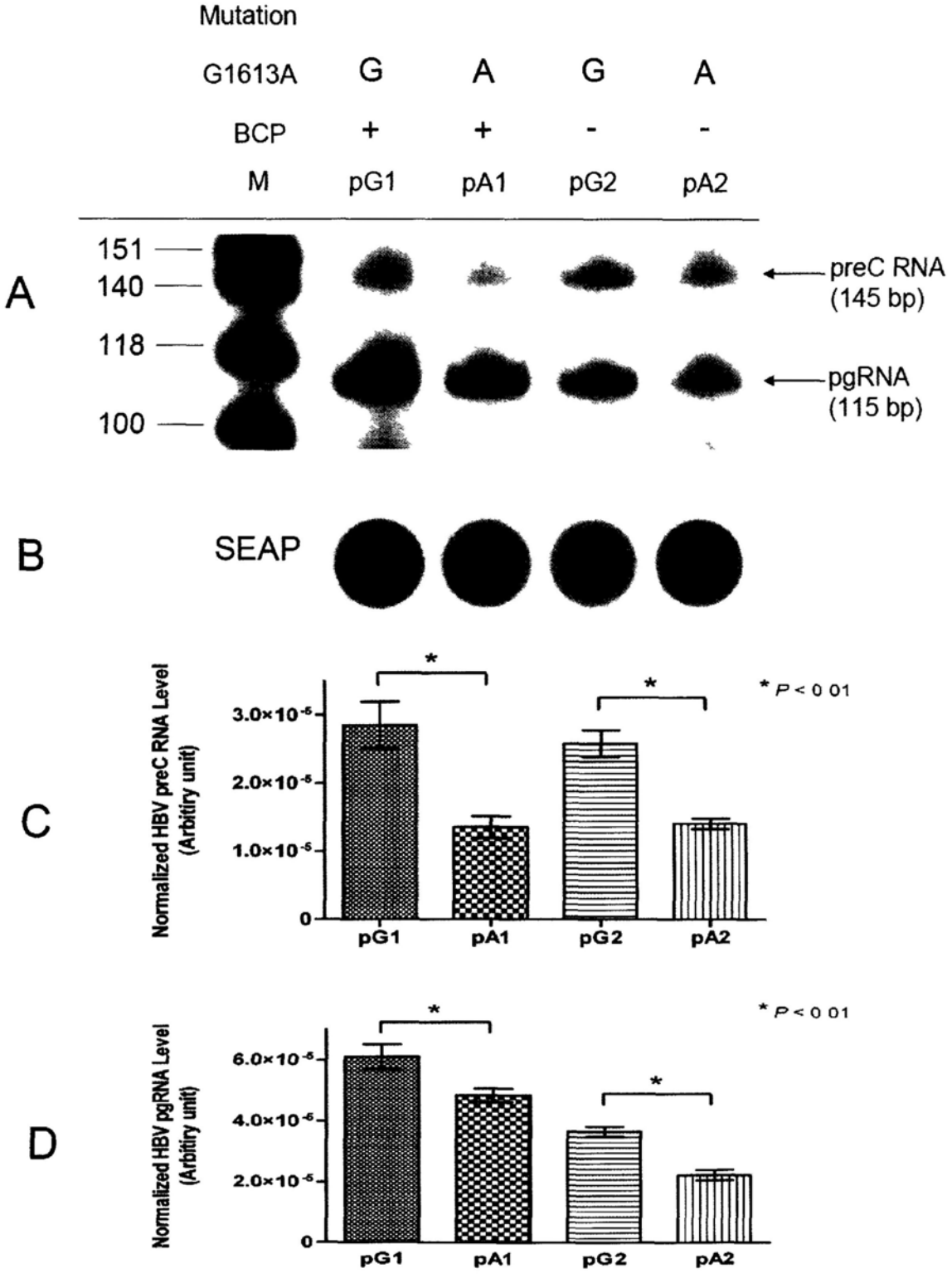

Mutation
G1613A G A G A
BCP + + - -
M pG1 pA1 pG2 pA2
A
151
140
118
100
preC RNA
(145 bp)
pgRNA
(115 bp)
B
SEAP
C
Normalized HBV preC RNA Level
(Arbitiry unit)
3.0×10⁻⁵
2.0×10⁻⁵
1.0×10⁻⁵
0
pG1 pA1 pG2 pA2
* P < 0 01
D
Normalized HBV pgRNA Level
(Arbitiry unit)
6.0×10⁻⁵
4.0×10⁻⁵
2.0×10⁻⁵
0
pG1 pA1 pG2 pA2
* P < 0 01

Fig. 3.11 Result of primer extension assay showed the effect of G1613A mutation on the relative amount of the preC mRNA and pgRNA transcription. Total RNA was extracted at 5-day posttransfection of 1.3-fold HBV genomes in HuH7. The constructs had the same DNA background except for the site of G1613A and BCP mutations as indicated in the figure. 10 μg of RNA was used in each primer extension reaction. The preC mRNA (145 bp) and pgRNA (115 bp) were successfully amplified and revealed on the 8% acrylamide gel. The gel photo is the representative from 4 independent experiments (A). The relative levels of the preC mRNA and pgRNA were normalized to the co-transfected SEAP activity (B). Results were presented as the mean values ± S.D. from 4 independent experiments for preC mRNA (C) and pgRNA (D).

3.6 Overexpression of Candidate Proteins that Bind on the HBV Negative Regulatory Element

By literature search, there are 2 protein candidates that have been reported to bind on the NRE on the HBV core promoter. They are the Negative Regulatory Element Binding Protein (NREBP) (Sun *et al.*, 2001; Lo and Ting, 1994) and the regulatory factor X1 (RFX1) (Buckwold *et al.*, 1997). The full-length NREBP cDNA is about 8 kb long and the predicted protein size is 262kDa, which is too large and not easily cloned. Therefore, only the DNA-binding domain of NREBP (BP15) had been cloned in our study.

3.6.1 Overexpression of myc-tagged BP15 and RFX1 proteins

The DNA-binding domain of NREBP (BP15) (nt. 3939-5639) was cloned into pCMV-myc vector with an N-terminal myc-tag, and the RFX1 protein expression construct with a C-terminal myc-tag was purchased from Origene (www.origene.com) (Cat no. RC207872). These two constructs were named pCMV-BP15 and pCMV-RFX1 respectively in our study. They were transiently transfected into human

hepatoma HuH7 cells for overexpression and examined by western blot using mouse monoclonal anti-myc antibody (Santa Cruz). Fig. 3.12A showed these proteins were overexpressed after transient transfection. The β-actin was the loading control of proteins (Fig. 3.12B).

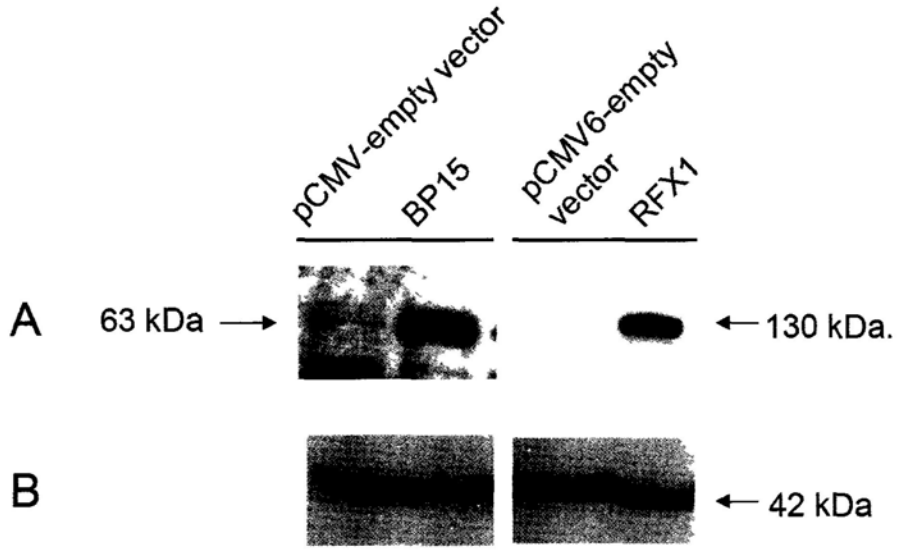


Fig. 3.12 Western blot showing the overexpression of myc-tagged BP15 (A) and RFX1 proteins and β-actin (B) at 2-day of transient transfection in HuH7 cells. 10 μg of protein was used in each lane. Primary antibody of mouse anti-myc (1:1000 in 3% BSA in TBST) was added and incubated at 4℃ overnight. After washing with TBST, secondary antibody of goat polyclonal anti-mouse conjugated with HRP was added (1:5000) at room temperature for 1 hour followed by washing with TBST and film exposure. Myc-tageged BP15 was successfully expressed as 63kDa, while myc-tagged RFX1 was a protein of 130kDa. The house keeping gene, β-actin of 42 kDa, was examined as the internal loading control.

3.6.2 Differential expression of endogenous NREBP in HCC tissues and non-tumorous counterparts

In order to investigate the involvement of NREBP in HBV-related HCC, we obtained cDNA of 16 samples of HCC tissues and their neighboring non-tumorous counterparts to examinate the endogenous NREBP expression. The cDNA of the tissues were provided by Prof. Winnie Yeo in Prince of Wales Hospital. The clinical background of the samples is listed in Table 3.5. All of the samples are HBV-positive. The transcriptional level of the endogenous NREBP in HCC tissues and their non-tumorous counterparts were examined by real-time quantitative PCR. Among the 16 pairs of samples, 11 of them (68.8%) showed significant higher transcription level of NREBP in the non-tumorous counterparts for more than 2 folds, when compared to their tumor counterparts (Fig. 3.13). In addition, when the samples were grouped according to the presence of hepatitis and cirrhosis, the average fold change of NREBP in samples of hepatitis group (4.48 folds) was higher than that of cirrhosis group (2.49 folds).

Tissue No.	Age	Sex	Diagnosis	Liver status	AFP level	HBV
360	45	F	HCC	**c**	32	p
370	43	M	HCC	**c**	48	p
381	59	M	HCC	**c**	121	p
382	58	M	HCC	**c**	8	p
383	43	M	HCC	**c**	7720	p
389	62	M	HCC	**c**	2	p
414	72	F	HCC	**c**	35	p
426	45	M	HCC	**c**	116600	p
368	39	M	HCC	**h**	2	p
385	52	M	HCC	**h**	4	p
412	27	M	HCC	**h**	380	p
418	59	M	HCC	**h**	480	p
420	59	M	HCC	**h**	4300	p
422	53	M	HCC	**h**	610	p
425	38	M	HCC	**h**	14200	p
427	53	M	HCC	**h**	3	p

HCC: hepatocellular carcinoma; Liver status: c= cirrhosis; h= hepatitis only, no cirrhosis; AFP: α-fetoprotein; HBV: p= positive; n= negative

Table 3.5 Table showing the clinical background of 16 HCC patients.

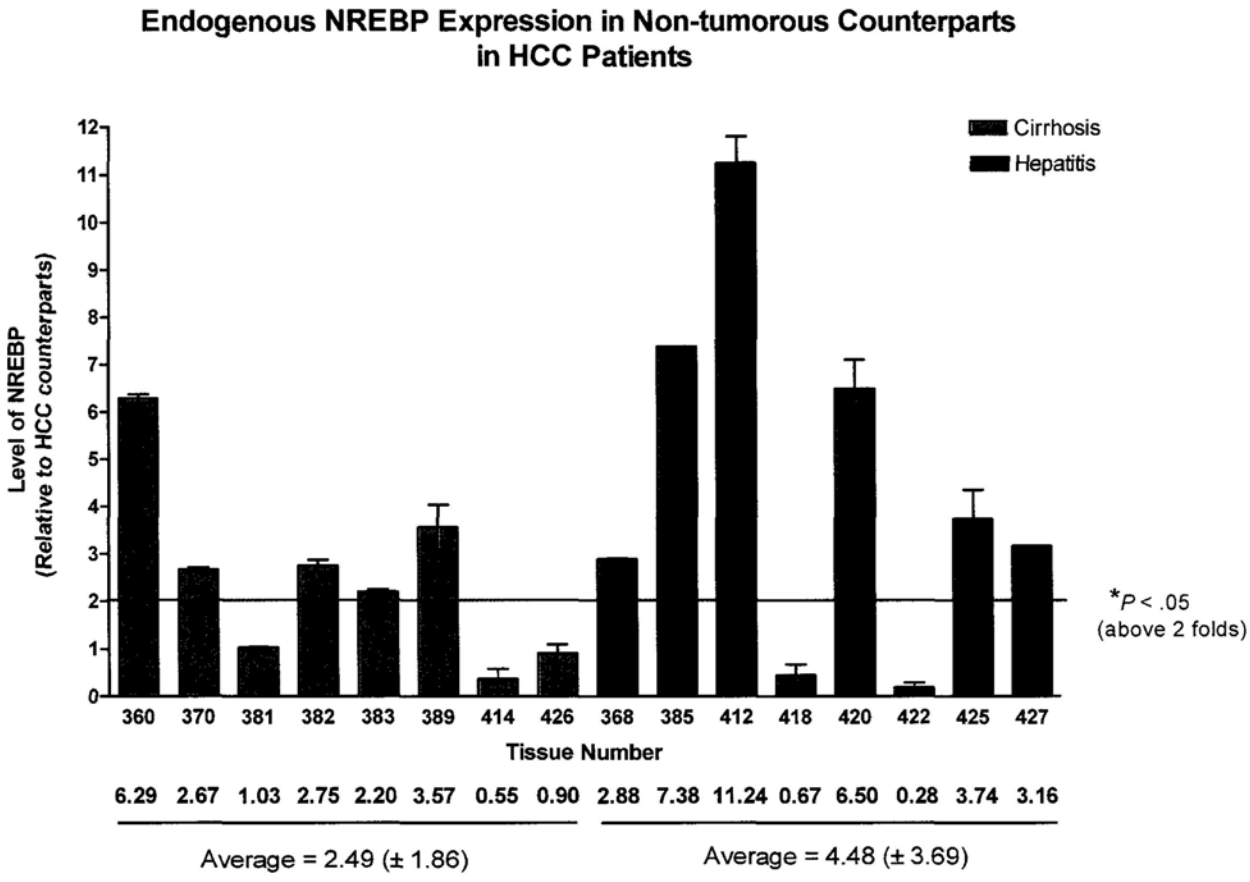


Fig. 3.13 The level of endogenous NREBP mRNA in non-tumorous counterparts relative to the HCC counterparts in 18 HCC liver samples. 11 out of 16 pairs of samples showed more than 2-folds of NREBP mRNA level in non-tumorous counterparts when compared to that in their HCC counterparts. Results were normalized to the housekeeping control, glyceraldehyde -3- phosphate dehydrogenase (GAPDH), mRNA level, and presented as relative fold change to the tumor counterparts (mean ± S. D.) in triplicates.

3.6.3 Co-transfection of protein-expression plasmids and HBV core promoters

3.6.3.1 BP15 overexpression and core promoter activity

To investigate whether the BP15 protein could exert its effect on the HBV core promoter and its A1613-mutant, the pCMV-BP15 construct was transiently co-transfected with pGL3-core promoters in HuH7 cells and examined by luciferase assay at 2-day posttransfection. As shown in Fig 3.14, the overexpression of BP15 protein slightly decreased the HBV core promoter activities by about 20%, in wild-type (G1613) and mutant (A1613) core promoter constructs respectively. The pGL3-Basic and HBV preS1 promoters served as controls. There was a slight decrease in the preS1 promoter activities by 20% in wild-type (T2712) and mutant (C2712) respectively, after the transient overexpression of the BP15 protein. This indicated that the overexpressed BP15 protein has no specific effect on the HBV core promoter activity. This may be due the BP15 protein is not the full-length functional protein NREBP. Therefore the specific effect of the DNA-binding domain BP15 on core promoter could not be revealed in this assay.

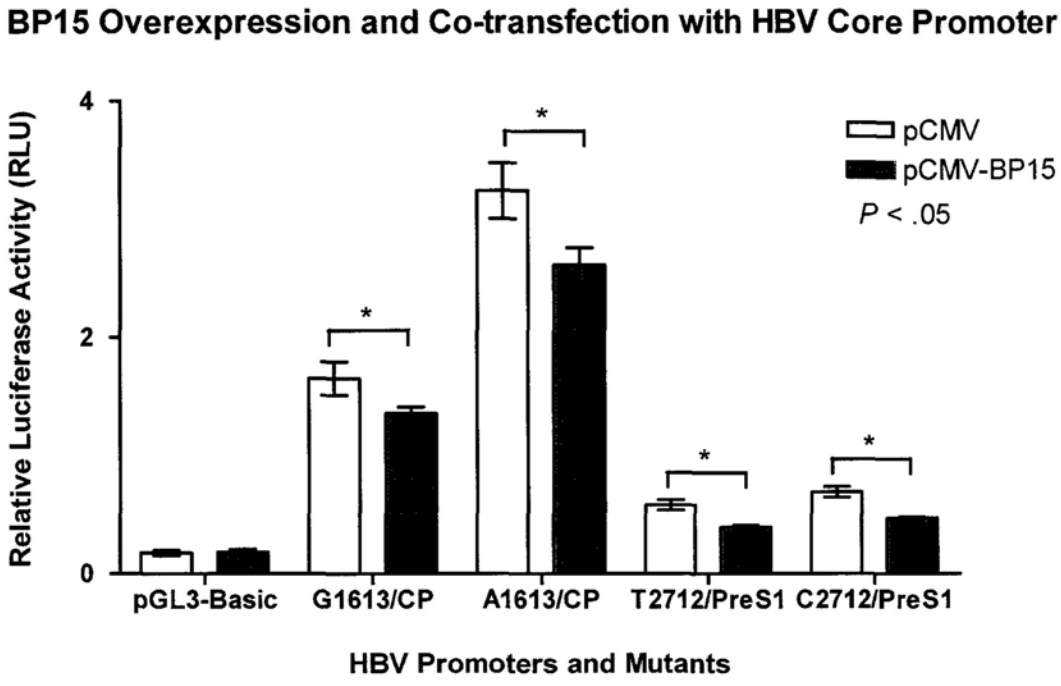


Fig. 3.14 The relative luciferase activity of the HBV promoters after the transient overexpression of BP15 protein at 2-day posttransfection in HuH7 cells. 0.1 μg of each pCMV-BP15 plasmid and pGL3-promoter plasmid and 0.001 μg of pCMV-PRL were transfected. The overexpression of BP15 caused a slight decrease in HBV core and preS1 promoter activities, in wild-type and mutant. Results were presented as relative luciferase activity in 3 independent experiments with triplicates (mean ± S. D.) after normalizing to the co-transfected *Renilla* luciferase activities. CP: HBV Core promoter; PreS1: HBV preS1 promoter.

3.6.3.2 RFX1 overexpression and core promoter activity

Similarly, to investigate whether the RFX1 protein could exert its effect on the HBV core promoter and its A1613-mutant, the pCMV-RFX1 construct was transiently co-transfected with pGL3-core promoters in HuH7 cells and examined by luciferase assay at 2-day posttransfection. Result showed that the overexpression of RFX1 protein drastically enhanced the HBV G1613 wild-type core promoter activity, by 2.7 folds and further increased the A1613 mutated core promoter by 1.8 folds (Fig. 3.15). In contrast, the activities of the controls, the pGL3-Basic and HBV preS1 promoters, did not show any significant change after the RFX1 protein expression. This indicated that the RFX1 protein could act on the core promoter specifically and induced its activity.

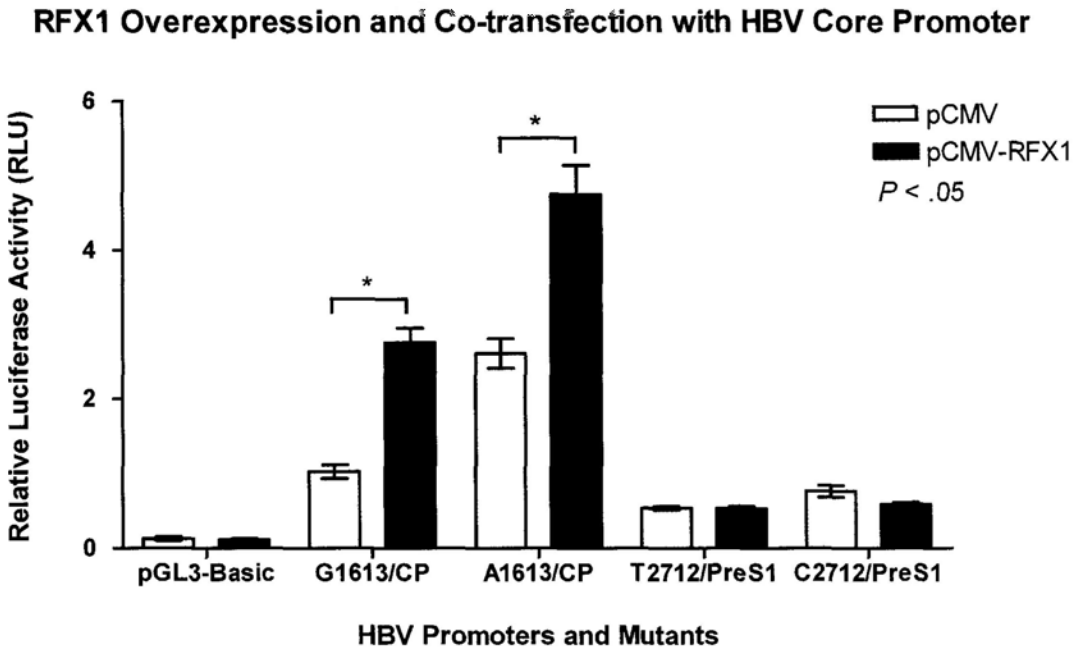


Fig. 3.15 The relative luciferase activity of the HBV promoters after the transient overexpression of RFX1 protein at 2-day posttransfection in HuH7 cells. 0.1 μg of each pCMV-RFX1 plasmid and pGL3-promoter plasmid and 0.001 μg of pCMV-PRL were transfected. The overexpression of RFX1 caused a significant increase in HBV core promoters in wild-type (G1613) and mutant (A1613) respectively. The control preS1 promoters did not response to the RFX1 overexpression. Results were presented as relative luciferase activity in 3 independent experiments with triplicates (mean ± S. D.) after normalizing to the co-transfected *Renilla* luciferase activities. CP: HBV Core promoter; PreS1: HBV preS1 promoter.

3.7 Electromobility Shift Assay

3.7.1 Endogenous proteins bind to wild-type and mutant NRE probes

To investigate if there are any endogenous proteins bind to the NRE region, Electromobility Shift Assay (EMSA) using HuH7 nuclear extract and ^{32}P end-labeled NRE wild-type and A1613 mutant probe were examined. The non-specific oligo (NS) was used as a control. Fig. 3.16 showed two protein complexes (named C1 and C2) in HuH7 nuclear extract that can shift the labeled probes, but not the NS control. Besides, the wild-type probe seemed to bind to C1 at a higher affinity when compared to the mutant probes. The cytosolic fraction served as a protein control in the assay and it did not cause any specific band shift in the binding neither with NRE oligo nor the NS control.

Fig. 3.17 showed the intensity of the band shifted was directly proportional to the amount of nuclear extract used in the binding reaction. Similar to the result shown in Fig. 3.16, C1 seemed to have differential binding ability between the wild-type and mutant NRE probes and bound to the wild-type probes with a stronger affinity. In

contrast, C2 seems to preferentially bind to the mutant probes more strongly than the wild-type probe. Moreover, to further investigate the difference in the binding ability between the wild-type and mutant probes in C1 and C2, unlabeled wild-type and mutant oligos in increasing concentrations were used to compete with the labeled probes. As shown in Fig. 3.18, the band caused by C1 was totally disappeared by adding 250X unlabeled wild-type oligos. In contrast, the band could not be competed when using 250X of the unlabeled mutant oligo. This indicated the C1 protein bound to wild-type probe with higher affinity than that to the mutant probe. On the other hand, the band caused by C2 was easily displaced by adding small amount of unlabeled competitors (e.g 10X). Furthermore, the addition of NS competitor could not abolish the bands contributed by C1 and C2, which reveals the binding in C1 and C2 to the NRE are specific.

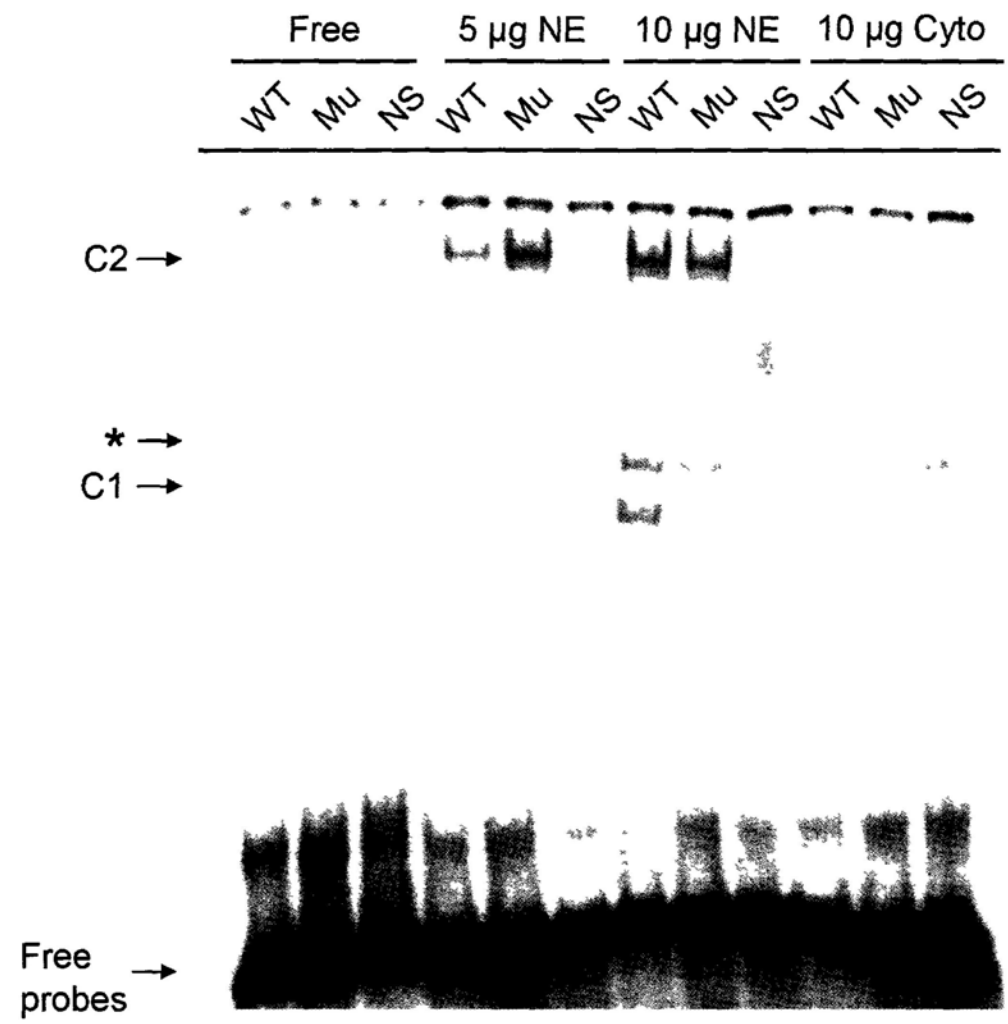


Fig. 3.16 Binding of HuH7 nuclear and cytosolic fractions to the wild-type and mutant NRE oligos. 5-10 µg of HuH7 nuclear extracts were used to bind to the G1613 wild-type and A1613 mutant NRE probe. Result showed there are two specific protein complexes (C1 and C2) bound to the NRE probes, with different binding ability between the wild-type and mutant NRE probes. The cytosolic fraction and non-specific probe controls (NS) showed no specific band shift in the gel. No protein was added in the free probe controls and also no band shift was observed. NE: Nuclear Extract; Cyto: cytosolic fraction; WT: Wild-type probe; Mu: Mutant probe; NS: Non-specific probe; C1: Protein complex I; C2: Protein complex II; *: non-specific band.

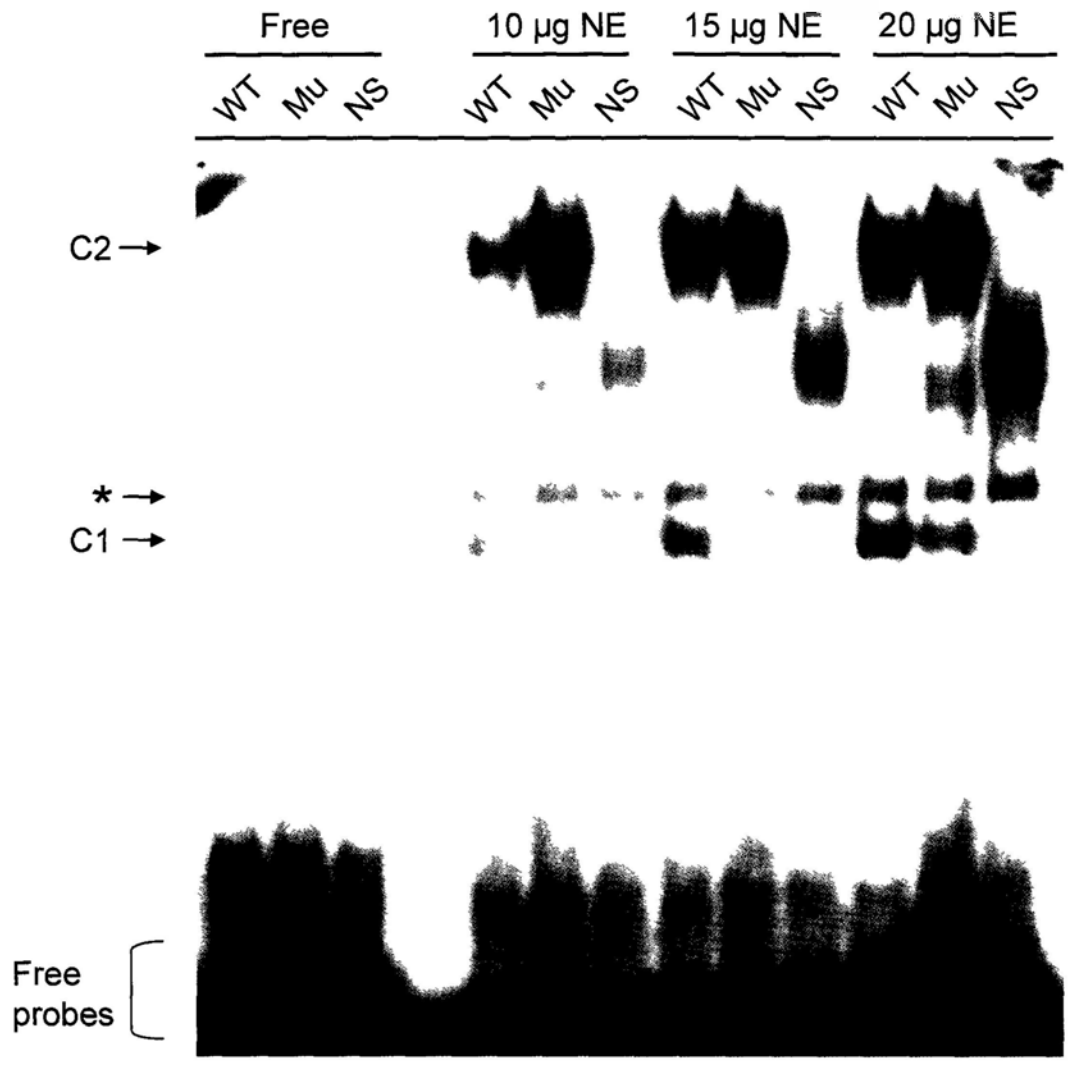


Fig. 3.17 Figure showing the titration of the HuH7 nuclear extract and binding to the wild-type and mutant NRE probes. 10-20 µg of protein were used to bind with G1613 wild-type, A1613 mutant and non-specific probes as indicated in the figure. As the amount of protein increased, the band intensity for the protein complex C1 and C2 increased proportionally. C1 has a stronger binding ability to the wild-type probe than that to the mutant probe, while C2 bound to the mutant probe stronger than the wild-type probe. No protein was added in the free probe controls and no band shift was observed. NE: Nuclear Extract; WT: Wild-type probe; Mu: Mutant probe; NS: Non-specific probe; C1: Protein complex I; C2: Protein complex II; *: non-specific band.

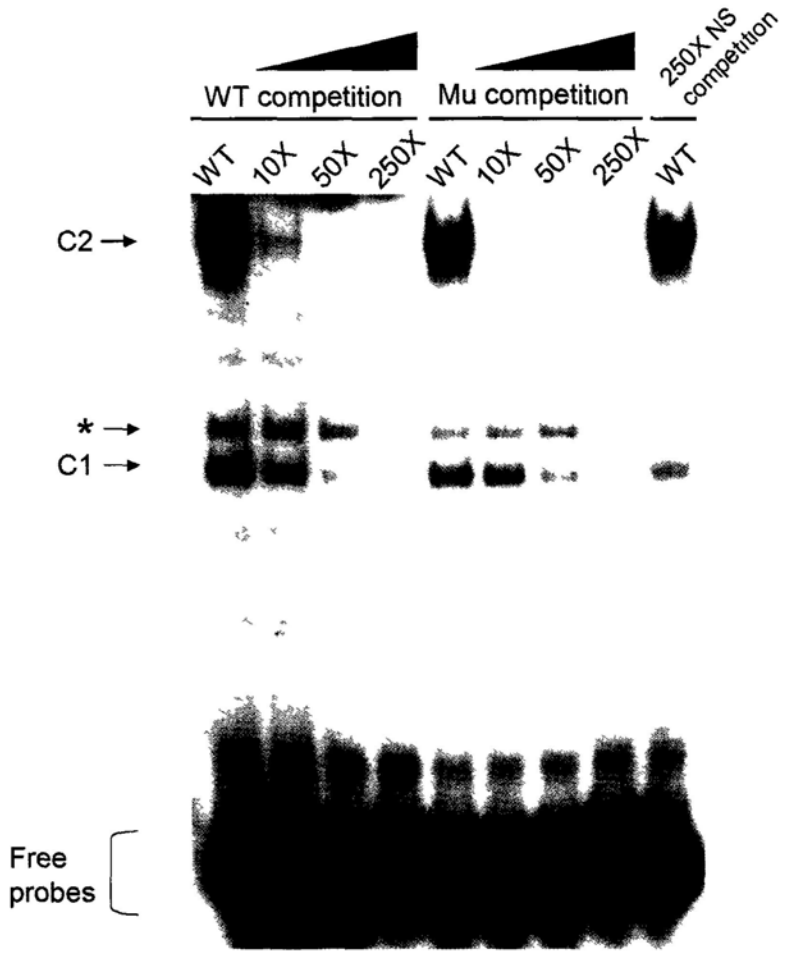


Fig. 3.18 The competition binding using unlabeled wild-type and mutant NRE oligos. 10 μg of HuH7 nuclear extract was used in each lane. Increasing concentration of the cold wild-type and mutant oligonucleotides were added to compete the specific binding between the probes and proteins. As the concentration of the unlabeled wild-type and mutant oligos increased, the band intensity of C1 and C2 decreased. The band of C1 was totally displaced under 250X of unlabeled wild-type but not the mutant oligo indicating C1 had higher affinity towards the wild-type oligonucleotides. In contrast, the band of C2 was easily competed by adding low concentration of unlabeled competitors indicating C2 had higher affinity towards the mutant oligonucleotide. NE: Nuclear Extract; WT: Wild-type probe; Mu: Mutant probe; NS: Non-specific probe; C1: Protein complex I; C2: Protein complex II; *: non-specific band.

3.7.2 *In vitro* transcribed/translated RFX1 and NRE oligo binding

The identities of proteins in C1 and C2 are unknown. By literature search, NREBP and RFX1 are the proteins likely to bind on the NRE region on core promoter. As the size of the NREBP is too large, *in vitro* translation of this protein is not feasible. Therefore, we have cloned the RFX1 coding sequence into another protein expression vector (pCMV-TNT) and expressed the protein using an *in vitro* protein synthesis system (Promega). Fig. 3.19 showed the RFX1 was successfully transcribed and translated *in vitro* using radioactive ^{35}S methionine. The smear of the bands may due to protein degradation or partially translated product for a large size of protein (apparent size is 130 kDa). The protein lysate obtained by using non-radioactive methionine in the same protocol was used in gel shift experiment. As shown in Fig. 3.20, the expressed RFX1 protein shifted the wild-type and mutant NRE probes to a position correspond to C2. This suggested that RFX1 may be one of the endogenous proteins in HuH7 that can bind to the NRE region on the HBV core promoter. Moreover, under the same amount of protein loading, the mutant NRE probe had a stronger binding ability to the RFX1, indicated by stronger band intensity. Besides, there was no protein shifted in the NS control, indicated the protein complex was specific to the NRE probes.

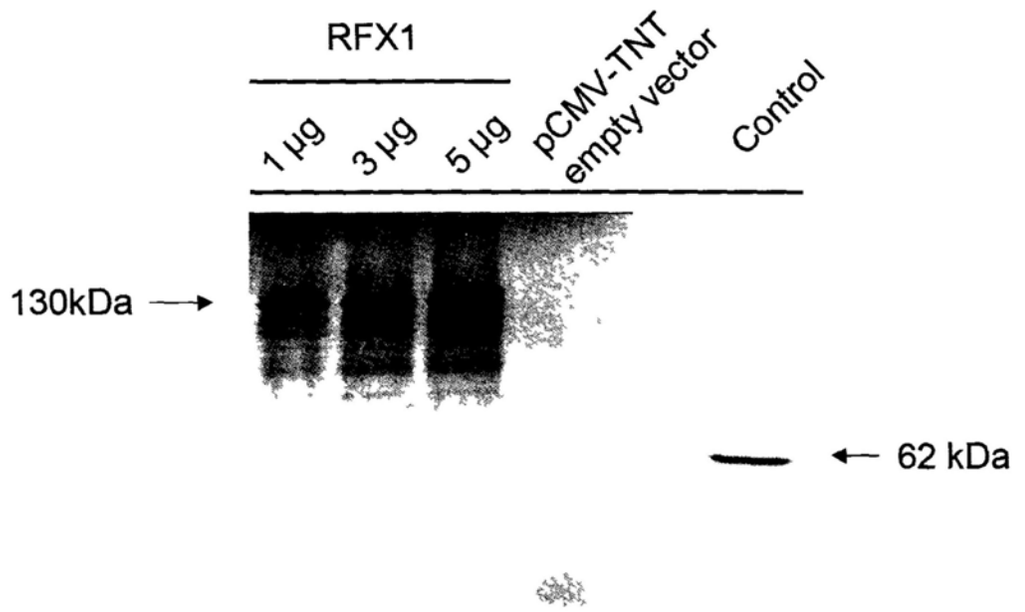


Fig. 3.19 *In vitro* expression of RFX1 using radioactive ^{35}S methionine. 1-5 µg of RFX1-expression construct was used for each reaction. The *in vitro* synthesized RFX1 protein was revealed using 10% SDS-PAGE and the signals were detected by autoradiography for 2 days. RFX1 of apparent size of 130 kDa was successfully expressed. The smear observed may be due to the degradation of proteins of large size. No band was shown in the empty vector control. The 62 kDa protein in the control correspond to the luciferase protein expressed from a control plasmid provided by the kit.

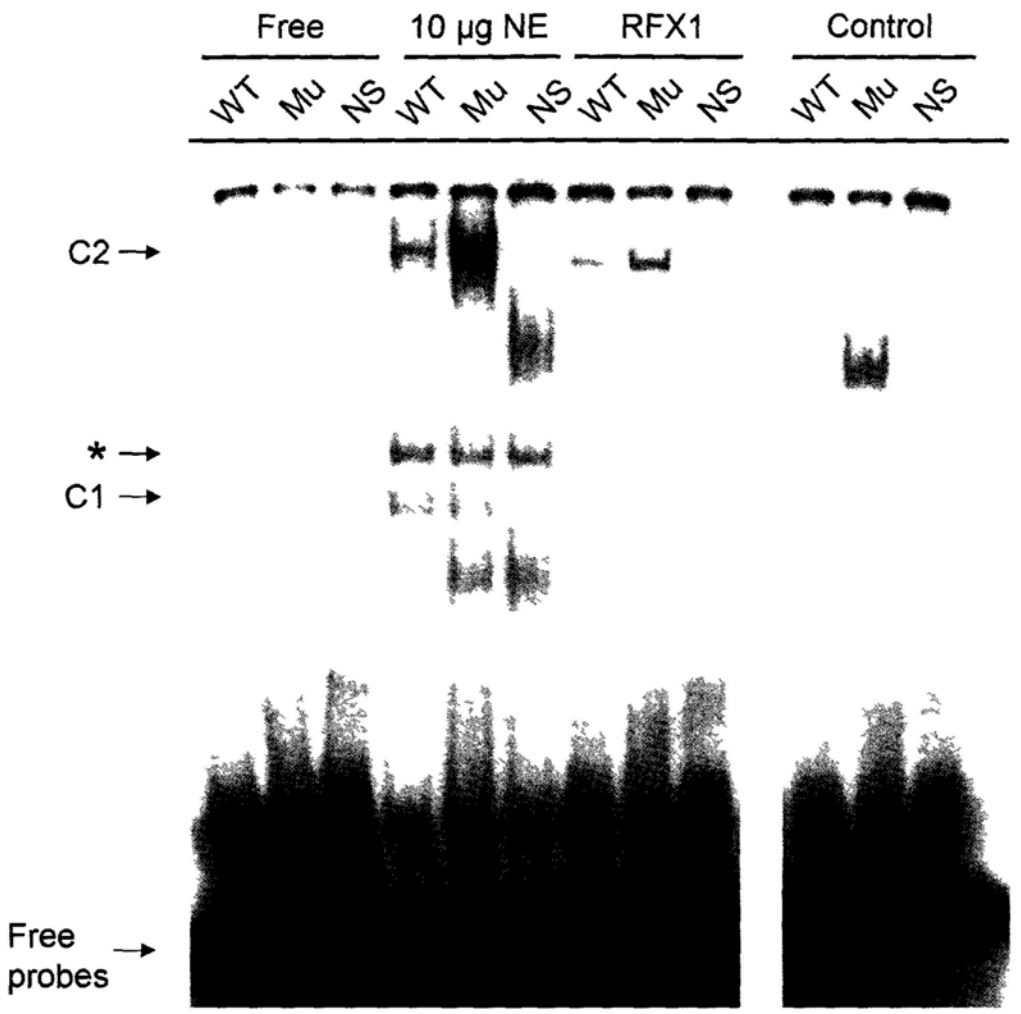


Fig. 3.20 Gel shift assay using *in vitro* translated RFX1 protein binding to wild-type and mutant NRE probes. 10 µg of HuH7 nuclear extract and 3 µl of *in vitro* synthesized RFX1 or control proteins were used in the binding. It showed the *in vitro* expressed RFX1 protein caused the shifting of the probes to a position correspond to C2. Besides, the RFX1 protein had a higher binding affinity to the mutant probe than the wild-type probe for a band of higher intensity, which was also observed in C2. RFX1 maybe one of the candidate protein that binds to the NRE region in HBV core promoter. NE: HuH7 Nuclear Extract; WT: Wild-type probe; Mu: Mutant probe; NS: Non-specific probe; C1: Protein complex I; C2: Protein complex II; *: non-specific band; Control: *in vitro* transcribed luciferase protein from the kit.

Chapter 4 DISCUSSIONS

4.1 The C1165T and T2712C Mutations do not Affect the Promoter Activity

In the beginning of our study, we aimed to investigate the individual effect of the HCC-related hot spot mutations in HBV promoters, which has been identified in our previous study (Sung *et al.*, 2008). Among the 3 mutations in our study, the C1165T mutation on X promoter and T2712C mutation on PreS1 promoter did not confer any change on their respective promoter activity. Despite of that, their potential effects on the overlapped HBV genes cannot be neglected. Because the ORFs in the genome of HBV is partially overlapping. Any single mutation on the genome may affect more than one gene/promoter depends on the position of the mutation.

The nucleotide 1165 is located on the X promoter overlapped to the enhancer I region, but it is also situated at the RNaseH domain of the polymerase gene. It caused an amino acid change from serine to proline at residue 2 of the polymerase. According to Lai and Locarnini (2002), the RNaseH domain is responsible for the degradation of RNA template after the minus-strand DNA synthesis during viral replication. However, this step is not critical when compared to the priming step and the reverse transcription in the life cycle of HBV. Until now, there is no report on mutations on RNaseH that can significantly alter the viral replication. In this regards, the potential effect of this mutation on the virus still needs further investigation.

For nucleotide 2712, it is not only located at PreS1 promoter, but also located at the polymerase gene of HBV. The T2712C mutation could lead to an amino acid change from tyrosine to histidine at residue 136, which is located at the terminal protein (TP) domain of the viral polymerase (tpY136T). The TP domain

mainly involves in the priming step during the initiation of minus-strand DNA synthesis from pgRNA (Lai and Locarnini, 2002). It was achieved by the formation of a phosphodiester bond between the tyrosine residue 63 and the first nucleotide of the minus-strand DNA (Lanford *et al.*, 1997). Therefore, the effect of the tpY136T on viral polymerase, especially on the priming step and viral replication capacity should not be overlooked.

4.2 Prevalence of G1613A Mutation in HCC and HBV Genotypes

There are some studies showed that the G1613A mutation was commonly found in HCC patients (Cheng *et al.*, 2006; Shinkai *et al.*, 2007; Takahashi *et al.*, 1998). However, the effect of the G1613A mutation on the virus has not been studied before. Our study demonstrated that, in genotype B and C, the prevalence of G1613A mutation increased in HBV-related HCC when compared to HBV carriers a case-control study. Moreover, the mutation was significantly associated to HCC in HBV subgenotype Cs. In correlation to the core promoter activity, we have demonstrated that the G1613A mutation can up-regulate the HBV core promoter activity in different genetic background including genotype B and C. In general, the induction in genotype C was stronger than that in genotype B. The result suggested that the effect of G1613A mutation could be genotype-dependent. Besides, despite of the mutation was not found associated to subgenotype Ce in our study, the induction of promoter activity was the highest. There are 2 possible reasons for this discrepancy. First, the sample size in subgenotype Ce was very

small (N=10). The statistics would be less reliable in dealing with small sample size. Moreover, subgenotype Cs and Ce share great similarity in their genome sequences. By the definition of genotyping, their genome sequences only differ to each other by 4-8%. Therefore, it is predictable that the mutation may have similar effect on the core promoter activity in subgenotypes Cs and Ce. In contrast, the genome sequence of genotype B differs from that of genotype C by at least 8%. This is likely to be the reason why the promoter activities were induced differently by the same mutation. Second, the template in the promoter study was selected from one individual in each genotype only. The relative promoter activities maybe sample-specific. Nevertheless, our study provided a good model to elucidate the relationship between the G1613A mutation and HBV genotypes/subgenotypes.

Besides, many reports showed that HBV carriers infected with different HBV genotypes/subgenotypes differs in clinical manifestations. Genotype C is associated with more aggressive liver disease, increased viral load and higher risk

to HCC when compared to genotype B (Kao *et al.*, 2000; Sumi *et al.*, 2003). However, the reasons remain poorly understood. Involvement of c-myc-regulated genes in HCC related to HBV genotype C (Iizuka *et al.*, 2006) may be one of the possible mechanisms leading to the genotypic response of the virus. In conclusion, our results suggested that studying the G1613A mutation on core promoter could help understand the genotypic effect of the virus.

4.3 Relationship Between G1613A Mutation and High Viral Load

Serum viral load is not only affected by viral factors, such as mutations on the viral genome, but also by host factors such as age, administration of antiviral drugs and immune response. Generally, the viral load is higher in HBeAg positive patients (Xie *et al.*, 2003). Persistent elevation of a serum HBV DNA level in the 10^4-10^7 range increases the risk of HCC (Liu *et al.*, 2008). Most recently, a study reported that high HBV viral load (more than 6 log copies/ml) is associated with HCC in both univariate and multivariate analysis (Wong *et al.*, 2010). In this study, we showed that the G1613A mutation was associated with a serum viral load of more than 6 log copies/ml in female carriers. The reason for the discrepancy between sexes is still not clear. This may be because the study was a univariate analysis. More clinical information is required for a multivariate analysis to be carried out.

In addition, using the *in vitro* full-length HBV genome study, we consistently demonstrated that the G1613A mutation significantly increased the extracellular HBV DNA level. This implied that the mutation enhanced the viral replication. In general, there are a few possible ways for the increase in the viral DNA: (i) increase the production of core protein; (ii) enhance the pgRNA encapsidation and (iii) enhance the virion packaging and secretion. Up to now, there are several core promoter mutations that have been identified to alter viral protein production or replication efficiency by creating additional transcription factor (HNF-1, HNF-3, HNF-4 and C/EBP) binding sites (Gunther *et al.*, 1996; Pult *et al.*, 1997; Zheng *et al.*, 2004). In our study we have demonstrated that the G1613A mutation on the NRE on HBV could play a role on changing the binding affinity of some endogenous nuclear proteins to the NRE region. Moreover, despite further investigations are needed, our result suggested that RFX1 protein could be one of those.

4.4 The G1613A Mutation Suppressed HBeAg Secretion through downregulating the preC mRNA transcription

HBeAg is known as an indicator of HBV infectivity and active virus replication in HBsAg-positive individuals (Cappel *et al.*, 1977, Nordenfelt *et al.*, 1975). It serves as an immunomodulatory protein to buffer the anti-core protein immune response in early viral infection. Once the host develops an anti-HBe immune response, seroconversion from HBeAg to anti-HBe usually occurs to prevent the destruction of HBV-infected hepatocytes through membrane-bound HBeAg. Therefore, as a result of selection pressure, viral variants with less or none HBeAg expression are thereby favoured for the viral survival (Parekh *et al.*, 2003; Tacke *et al.*, 2004). There has been considerable work on the possibility that some of the more rapidly progressing HBV infections result from exposure to HBeAg-negative mutants. It is also hypothesized that HBeAg functions to suppress cell-mediated immunity to the virus core protein by depleting the respective Th1 helper cells involved in stimulation of the core protein specific

CTL response. However, it fails to explain how viruses lacking the capacity to produce HBeAg often replace wild-type viruses, as an infection evolves. Nonetheless, in the absence of experimental models to compare the intrinsic pathogenicity of the wild-type and mutant HBV strains, this issue is likely to remain unresolved.

The precore protein is a precursor protein of HBeAg, and it has been reported to inhibit viral replication *in vivo* in transgenic mice (Guidotti *et al.*, 1996). Our study demonstrated that the G1613A mutation is likely to favor for the viral production by diminishing the HBeAg secretion. And we revealed that the G1613A mutation suppressed HBeAg production on the transcription level of preC mRNA (Fig. 3.11C). The pgRNA is the template of viral DNA synthesis. Although we demonstrated that the G1613A mutation enhanced HBV DNA production (Fig. 3.9B), our result showed that the levels of pgRNA in the A1613 mutants were decreased. This reflected that there are some other factors played a

role on regulating the viral DNA production. In fact, HBV replication is a highly regulated process that involves a few critical steps, such as the encapsidation of the pgRNA into nucleocapsids, which requires a defined capsid environment and specific interaction between the encapsidated pgRNA and core particles. Also, functional viral polymerase and some heat-shock proteins are needed to be packaged together with the pgRNA in order for the reverse transcription of the pgRNA to viral DNA take place. Baumert *et al.* demostrated that the increase in pgRNA encapsidation and viral replication was largely independent of the level of pregenomic RNA transcription (Baumert *et al.*, 1998). In addition, some studies showed that the precore protein is inhibitory to viral replication *in vitro* and *in vivo* (Lambert, 1993; Scaglioni *et al.*, 1997 Guidotti *et al.*, 1996), and precore protein was shown to suppress HBV virus production at the encapsidation step (Guidotti et al., 1996; Lamberts et al., 1993, Buckwold et al., 1996). In this regards, it seems that the level of precore protein is a more critical factor in regulating the HBV replication when compared to the level of pgRNA

transcription. Taken together, our study indicated that G1613A mutation on the core promoter enhanced HBV replication and suppressed HBeAg secretion by down-regulating the preC mRNA transcription in the cells.

In contrast to HBeAg, we did not find any correlation of the HBsAg level to the G1613A mutation. HBsAg is an indicator of chronic viral hepatitis. It is a viral envelope protein on the surface of HBV. Therefore, it is expected that it will accompany increase in viral load. However, a correlation between the serum level of HBsAg and the HBV DNA level has not been found so far. In a few studies, the HBV DNA level was found to be higher in patients with a higher HBsAg serum level (Deguchi *et al.*, 2004, Lei *et al.*, 2006). However, there were patients with high HBV DNA level who tested negative for HBsAg (Mendenhall *et al.*, 1991; Pares *et al.*, 1990; Sawada *et al.*, 1993). In fact, the HBsAg are not only incorporated into virion envelopes but also appear as subviral particles without the component of core capsid and viral DNA secreted from the infected cells. These

subviral particles are usually highly overexpressed relative to virions. It is assumed that the massive HBsAg overproduction influences the host's immune system in a way that is advantageous for the virus.

4.5 Independent Effect of G1613A and BCP Mutation

ince the effect of the BCP mutation has been extensively characterized. It was included as the control in the luciferase assay. The effects of G1613A and BCP mutation are antagonistic. In Fig. 3.6, G1613A enhanced the promoter activity in regardless to the presence of the BCP mutation. In addition, the promoter activity of pC2 is higher than pC3, showing that the effect of G1613A counteracts the effect of BCP mutation, and could even further increases the promoter activity. This indicates that the G1613A mutation alters the core promoter activity through a way independent to that of the BCP mutation.

The BCP mutation has been extensively reported to decrease the HBeAg secretion and increase viral replication (Buckwold *et al.*, 1996; Parekh *et al.*, 2003; Tacke *et al.*, 2004). In our study, we demonstrated that the BCP mutation only caused a slight change in the core promoter activity (Fig. 3.6) together with an insignificant change in the HBeAg production (Fig. 3.8B) in subgenotype Cs.

These results are contradictory to the other studies. This is possibly caused by the use of HBV genome of different genotype background. In Fig. 3.6 when comparing pC1 and pC3, the decrease of promoter activity attributed by BCP mutation was retained in HBV subgenotype Cs. But surprisingly, this decrease is smaller than that we have obtained earlier using clones of genotype A (Fig. 3.3), with only 16.5%. Similar result was obtained when comparing pC2 and pC4.

In fact, the BCP mutation is highly prevalent in genotype C, indicating the BCP mutation is commonly found in both chronic carriers and HCC patients. This is also observed in our patient database. Dong *et al.* showed that the BCP mutation decreased the core promoter activity by about 50% in genotype B (Dong *et al.*, 2008). In contrast, Jammeh *et al.* (2008) showed the BCP mutation had no significant effect on the replication capacity of the virus in genotype C, which the result coincides with ours that the BCP mutation did not show significant effect on the intracellular and extracellular level of HBV DNA in the background of

subgenotype Cs. It suggested the BCP mutation could also exert its effect on HBV in genotype-dependent manner. This would be an interesting topic for investigation as the BCP mutation is commonly found in HBV-related HCC patients.

4.6 G1613A Mutation in Negative Regulatory Element of Core Promoter

4.6.1 G1613A is the only hotspot in NRE

We demonstrated the drastic effect of the G1613A mutation on enhancing the HBV core promoter activity. The effect was reversed when the nucleotide was back mutated from A to G. The core promoter can be divided into BCP and the CURS (Kramvis and Kew, 1999). CURS is a cis-acting element upstream of BCP and known to activate BCP in an orientation- and position-dependent manner (Yuh *et al.*, 1992; Guo *et al.*, 1993). Its stimulatory effect can be partially abolished by a regulatory sequence further upstream, known as negative regulatory element (NRE) (nt. 1613-1636). The NRE can repress the activity of CURS region of core promoter and enhancer II by 10-20 folds (Lo and Ting, 1994; Chen and Ou, 1995). The G1613A mutation is located at this region.

The NRE sequence from nucleotide 1604 to 1610 is not required for its negative regulatory function (Lo and Ting, 1994). A number of studies showed that mutations in region within nt. 1611-1619 destroyed most of the NRE activity and increased the core promoter activity (Chen and Ou, 1995; Gerlach *et al.*, 1992; Lo and Ting, 1994; Park *et al.*, 1997; Sun *et al.*, 2001). However, all these studies have not investigated the specific effects of the G-to-A mutation at nt. 1613. Moreover, in an alignment analysis of a total of 803 sequences of HBV full length genome obtained from NCBI database (Table 3.1), the G-to-A mutation at the nucleotide 1613 is the only hot spot for mutation within the NRE region (nt. 1611-1619). Here, we are the first group to demonstrate that a single mutation in the NRE can modulate the characteristics of the HBV.

4.6.2 The negative regulatory element binding protein (NREBP) and HBV core promoter

A protein that binds to the NRE called NREBP has been characterized (Sun et al, 2001). NREBP mRNA is 8,266 nucleotides in size and encodes a protein of 2386 amino acids (262 kDa). It is expressed in all tissues including liver and has shown to be localized in the nucleus in HuH7 cells (Sun *et al.*, 2001). Sun *et al.* (2001) also demonstrated that transient overexpression of a N-terminally truncated NREBP strongly repressed the HBV core promoter activity in the presence of NRE, resulting in the inhibition of the transcription of HBV genes and the production of HBV virions. Moreover, they showed that the middle portion of NREBP, called BP15 (amino acid 1298-1236), was the DNA binding fragment and can interact with wild-type NRE sequence (nt. 1611-1634) specifically but not with mutant NRE (contained mutations from nt. 1611-1619) by southwestern analysis. In regard to this, we have cloned the BP15 fragment and investigate its effect on HBV core promoter.

On the other hand, the consensus sequence for BP15 binding is AG(G/T)AN(C/G)(A/G)CC (Sun *et al.*, 2001). As shown in Fig. 4.1, this sequence perfectly matches with the NRE sequence in the HBV genome (GAGACCACC), in which the underlined nucleotide is corresponded to the nucleotide 1613 in our study. Although there are studies showed the NREBP can bind to the wild-type but not randomly mutated NRE sequence, the single mutation at nucleotide 1613 has not been tested before. We speculated that the G1613A mutation on the NRE sequence could weaken the binding of NREBP to the core promoter, which leads to the increase in the promoter activity.

However, as shown in Fig. 3.14, overexpression of BP15 could only cause a mild repression of the co-transfected core promoter, and the effect was not specific to core promoter. The possible reason would be the BP15 is the DNA-binding domain but it does not contain the activity domain that responsible for the function of the protein. Despite of this, the gel shift experiment revealed

that an endogenous protein of a smaller size form a protein complex (C1). We demonstrated that this complex would bind to wild-type NRE probe in a manner stronger than the mutant NRE, which contains the G1613A mutation, by competition experiment. Although the BP15 is not the full-length form of NREBP, it is naturally occurring inside hepatoma cells as it was initially isolated and identified by cDNA expression cloning using concatamers of NRE probe (Sun *et al.*, 2001). Furthermore, Fig. 4.1 showed that the G1613A mutation is located at the consensus BP15 binding site to the NRE region. Taken together, we cannot rule out the possibility that the protein in C1 complex contains the endogenous BP15 protein. However, at present the antibody against BP15 is not available, therefore the identity of the C1 protein complex remains to be identified.

In addition, we showed that the endogenous NREBP mRNA transcription level in tumor counterparts was lower than that in non-tumor counterparts in clinical samples. This suggested the possibility that the lower expression level of NREBP in tumor enhances the activity of HBV core promoter in HCC patients.

4.6.3 RFX1 increases the HBV core promoter activity possibly via the binding to NRE

The NRE contains three subregions which act synergistically to suppress core promoter activity. One of these subregions, called NREγ (nt. 1605-1625), is active in HuH7 cells. In contrast to NREBP, RFX1 is reported to be the transcription activator which binds to the NREγ on the core promoter (Buckwold *et al.*, 1997). The consensus RFX1 binding sequence has extensive homology to the sequence of NREγ, from nt. 1605 to 1617 (Reinhold *et al.*, 1995, Buckwold *et al.*, 1997). As shown on Fig. 4.1, the G1613A mutation on the NREγ site further matches the consensus RFX1 binding sequence. In addition, overexpression of RFX1 can activate the transcription in an NREγ-dependent manner (Chen and Ou, 1995). Taken together, we speculated that G1613A could enhance the binding of RFX1 and hence core promoter activity.

RFX1 is approximately 130 kDa in size. It was found to be the methylation-dependent DNA binding protein (MDBP) (Zhang *et al.*, 1993). Our

study showed that RFX1 overexpression could significantly enhance the HBV core promoter activity by 2-3 folds in wild-type and mutant (Fig. 3.15). This is consistent to a study demonstrating that the presence of NREγ can lower the core promoter activity by 2-3 folds in HuH7 cells (Chen and Ou, 1995). Moreover, we showed that the *in vitro* expressed RFX1 could shift the labeled NRE probes to the same position as that of the C2 complex in the gel shift experiment using the HuH7 nuclear extract (Fig. 3.20). A previous study had demonstrated that RFX1 synthesized *in vitro* could bind to the NREγ probe but not to the mutated NREγ probe (Buckwold *et al.*, 1997). Here we showed that RFX1 could bind to the mutant NRE probe at a higher affinity when compared to the wild-type, revealing that the G1613A mutation on NRE could favor the binding to RFX1. Together with the result of luciferase assay, it is possible that RFX1 could be the endogenous protein which can bind to the NRE regions, and enhanced the core promoter with G1613A mutation. However, since we do not have an antibody against RFX1, the identity of the protein complexes remains to be elucidated.

In addition, it has been proposed that RFX1 can form a heterodimer with the myc intron-binding protein (MIBP1) and possesses a gene-suppressive activity (Reinhold *et al* , 1995). Therefore, it is hypothesized that RFX1 functions as a positive regulator of the core promoter, and converted into a negative binding regulator when binds to MIBP1 to suppress core promoter through the NREγ site. This is suggested that the dual activity of RFX1 could be important for differential regulation of HBV gene expression during the viral life cycle (Buckwold *et al* , 1997; Reinhold *et al.*, 1995).

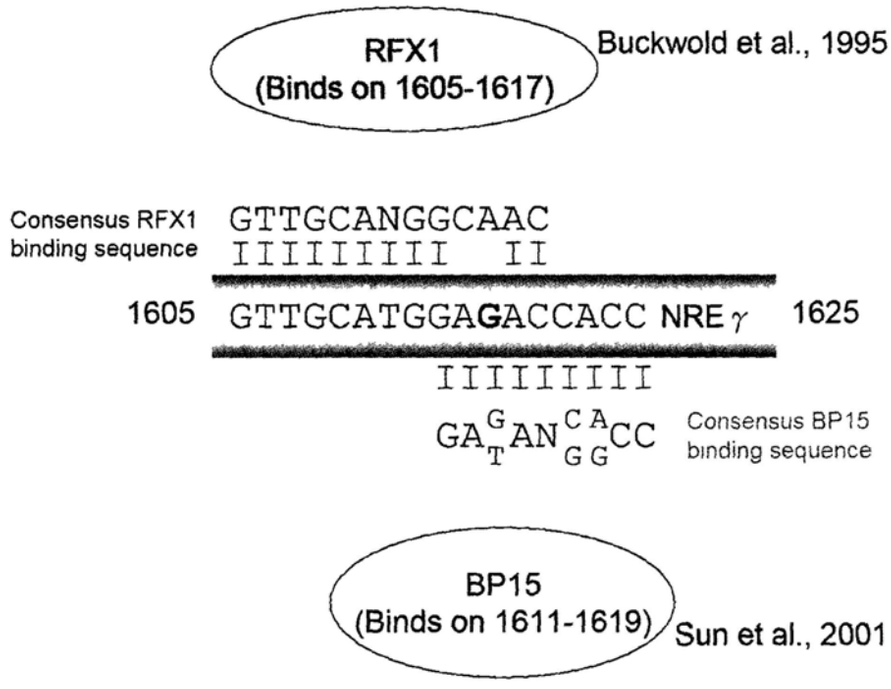


Fig. 4.1 Schematic diagram showing the consensus binding sequence of RFX1 and BP15 and the similarity to the NREγ site (nt. 1605-1625). The 1613 position on the NREγ site is indicated as the red 'G'. The A1613 NRE sequence perfectly matches the consensus binding sequence of RFX1, while the G1613 wildtype NRE sequence matches to the BP15 binding sequence.

4.7 Other Possible Effects of the G1613A Mutation on HBV

4.7.1 Effect on HBV core protein synthesis

Despite the effect of G1613A mutation on viral DNA, HBeAg and HBsAg that has been demonstrated, our attempts to reveal the core protein expression by Western blot analysis failed. One of the possible reasons could be the low transfection efficiency of the HBV genome in HuH7. Transient transfection of a GFP-tagged plasmid into HuH7 cells by the same protocol revealed that the transfection efficiency was only about 10-20%. Although it is considered as high enough for us to study the HBV replication (Lai and Locarnini, 2008), the level of protein expression may not be adequate for a conventional western blot analysis. In the contrary, the method used for quantification of HBeAg and HBsAg levels is much more sensitive. Another possible reason is that the primary antibodies used in our study were not specific enough to the target protein (HBcAg), although we had used more than one antibodies purchased from different companies.

4.7.2 Effect on HBV polymerase

Furthermore, the G1613A mutation is not only positioned on the core promoter, but also on the overlapped viral polymerase gene, leading to a amino acid change from arginine to lysine (R151K) on the RNaseH domain on the polymerase protein. As mentioned before, the RNaseH domain is responsible for the degradation of RNA template after the minus-strand DNA synthesis during viral replication (Lai and Locarnini *et al.*, 2002). Despite this step is not critical when compared to the priming step and the reverse transcription of HBV, we could not rule out the potential effect of G1613A mutation on the viral polymerase.

4.8 *In Vitro* Model System Used for Studying HBV Replication

The molecular study of HBV has advanced greatly during the last few decades, and results generated from a variety of model systems have deepen our understanding of the HBV life cycle. Unlike the other viruses, higher primates, such as chimpanzee, are the only animal models for HBV infection. Although *in vivo* models using related hepadnaviruses such as duck hepatitis B virus (DHBV) and woodchuck hepatitis virus (WHV) have been developed, the hosts for these viruses are distinct from HBV. Therefore, the importance of *in vitro* models for direct study of HBV has been emphasized. In our study, we have adopted an *in vitro* system to study HBV replication and protein production in HuH7 cell lines.

The advantages of using an *in vitro* system includes the availability of established cell lines and relatively inexpensive of cell culture system. Manipulation of cultured cells is easier and allow more experimental throughput than whole animals. However, the lack of convenient and reproducible cell culture

systems that support HBV infection has been a major impediment to the study of HBV. Fortunately, transfection of HBV full-length DNA makes it possible to study most aspects of HBV replication *in vitro*. Transient transfection is commonly employed to study HBV gene expression *in vitro* as it is rapid and flexible. A number of HBV mutants could be tested systematically without spending a long time to establish stable cell lines with mutant HBV genomes. However, transient transfection is lack of reproducibility and has a relatively lower level of HBV replication.

For optimal *in vitro* expression of HBV, the organization of HBV sequences in the transfected DNA construct should be considered. During a natural infection, the transcriptional template for HBV is the 3.2 kb cccDNA. The circular nature of this DNA allows production of the 3.5 kb pgRNA and preC mRNA. In regard to this, recombinant plasmid vectors containing a single unit-length HBV genome are thus not replication competent because at least one HBV coding region will be

interrupted. To overcome this, our study used HBV constructs of more than a genome length of DNA including the polyadenylation signal located downstream of the core promoter. Moreover, in well-differentiated hepatic cell lines such as HuH7, natural HBV promoters and enhancers function sufficiently to drive HBV transcription. In our study, we aim to investigate the effect of mutations on HBV promoter regions. Therefore, foreign promoter sequences such as cytomegalovirus (CMV) immediate early promoter could not be used to drive HBV expression because this would alter the ratio of HBV transcripts and trigger unknown effects on HBV replication. In regard to this, a cloning vector, pUC18, without mammalian expression promoter was used in our study.

HuH7 and HepG2 cells used here is suitable for HBV transfection study because they could support the viral replication. Both cell lines have been extensively characterized and are considered to be well-differentiated. Neither of them contains any integrated HBV DNA and therefore, HBV expression can be

attributed solely to transfected DNA.

Among the methods of transfection available, lipid-based transfection reagent, such as FuGENE6, was used in our study. Although the transfection efficiency is around 10-30%, it is considered to be high and enough to investigate the viral replication. When compared to the method using lentiviral or adenoviral vectors, which is of very high efficiency, such vectors have major drawbacks such as biosafety considerations, the expression of foreign proteins and toxicity.

4.9 Conclusion

To conclude, mutations C1165T and T2712C did not alter the X and preS1 promoters respectively. In contrast, by establishing a replication competent *in vitro* system to study HBV, we demonstrated that the G1613A mutation markedly decreased the extracellular HBeAg production by downregulating the preC mRNA transcription and lead to increased viral production. This was possibly through increasing the core promoter activity by regulating the binding of some endogenous proteins on the NRE of HBV. The identities of these proteins deserved further investigation. Despite of this, NREBP and RFX1 could be one of those. In addition, we demonstrated that the differential expression of NREBP in HCC and non-tumorous tissues could possibly play a role on HBV-related HCC. As a whole, our study was the first to demonstrate a novel single mutation on the NRE region could alter the viral replication and protein synthesis.

4.10 Future Perspectives

There are still a few ways through which the mutation G1613A could take part to alter the HBV characteristics via the core promoter, including the core protein synthesis and viral polymerase. They may also take a role in the viral pathogenesis which involves the process of pgRNA encapsidation and DNA replication. Together with the other two mutations in this study, the S2P mutation on the RNaseH domain (C1165T mutation) and tpY136T mutation on viral polymerase (T2712C mutation), these mutations await further investigation so as to broaden our understanding the molecular phathogenesis of HBV-related HCC.

Appendix

Alignment of sequences from chronic carriers with the HBV genotype B consensus sequence in carriers (B). The sequence from patient number 1939 is most similar to the consensus sequence. It was chosen as the template for X promoter (nt. 980-1373).

```
 Nt980                                                                1037
   B  GGAAAGTATGTCAACGAATTGTGGGTCTTTTGGGGTTTGCCGCCCCTTTCACGCAATG
1939  GGAAAGTATGTCAACGAATTGTGGGTCTTTTGGGGTTTGCCGCCCCTTTCACGCAATG

Nt1038                                                               1095
   B  TGGATATCCTGCTTTAATGCCTTTATATGCATGTATACAAGCRAAACAGGCTTTTACT
1939  TGGATATCCTGCTTTAATGCCTTTATATGCATGTATACAAGCAAAACAGGCTTTTACT

Nt1096                                                               1153
   B  TTCTCGCCAACTTACAAGGCCTTTCTRMGTAAMCAGTATCTGAACCTTTACCCCGTTG
1939  TTCTCGCCAACTTACAAGGCCTTTCTAAGTAAACAGTATCTGAACCTTTACCCCGTTG

Nt1154                                                               1211
   B  CTCGGCAACGGCCTGGTCTGTGCCAAGTGTTTGCTGACGCAACCCCCACTGGTTGGGG
1939  CTCGGCAACGGCCTGGTCTGTGCCAAGTGTTTGCTGACGCAACCCCCACTGGTTGGGG

Nt1212                                                               1269
   B  CTTGGCCATAGGCCATCAGCGCATGCGTGGAACCTTTGTGTCTCCTCTGCCGATCCAT
1939  CTTGGCCATAGGCCATCAGCGCATGCGTGGAACCTTTGTGTCGCCTCTGCCGATCCAT

Nt1270                                                               1327
   B  ACTGCGGAACTCCTAGCCGCTTGTTTTGCTCGCAGCAGGTCTGGRGCAAAACTCATCG
1939  ACTGCGGAACTCCTAGCCGCTTGTTTTGCTCGCAGCAGGTCTGGGGCAAAACTCATCG

Nt1328                                                    1373
   B  GGACTGACAATTCTGTCGTGCTCTCCCGCAAGTATACATCMTTTCC
1939  GGACTGACAATTCTGTCGTGCTCTCCCGCAAGTATACATCATTTCC
```

Alignment of sequences from HCC patients with the HBV genotype B consensus sequence in HCC (BHCC). The sequence from patient number 14 is most similar to the consensus sequence. It was chosen as the template for X promoter (nt. 980-1373).

```
  Nt980                                                                1037
BHCC  GGAAAGTATGTCAACGAATTGTGGGTCTTTTGGGGTTTGCCGCCCCTTTCACGCAATG
  14  GGAAAGTATGTCAACGAATTGTGGGTCTTTTGGGGTTTGCCGCCCCTTTCACACAATG

 Nt1038                                                               1095
BHCC  TGGATATCCTGCTTTAATGCCTTTATATGCATGTATACAAGCRAAACAGGCTTTTACT
  14  TGGATATCCTGCTTTAATGCCTTTATATGCATGTATACAAGCGAAACAGGCTTTTACT

 Nt1096                                                               1153
BHCC  TTCTCGCCAACTTACAAGGCCTTTCTRMGTMAACAGTATCTGAACCTTTACCCCGTTG
  14  TTCTCGCCAACCTACAAGTCCTTTCTAAGTAAACAGTATCTGAACCTTTACCCCGTTG

 Nt1154                                                               1211
BHCC  CTCGGCAACGGCCTGGTCTGTGCCAAGTGTTTGCTGACGCAACCCCCACTGGTTGGGG
  14  CTCGGCAACGGTCTGGTCTGTGCCAAGTGTTTGCTGACGCAACCCCCACTGGTTGGGG

 Nt1212                                                               1269
BHCC  CTTGGCCATAGGCCATCAGCGCATGCGTGGRACCTTTGTGTCTCCTCTGCCGATCCAT
  14  CTTGGCCATAGGCCATCAGCGCATGCGTGGAACCTTTGTGTCTCCTCTGCCGATCCAT

 Nt1270                                                               1327
BHCC  ACTGCGGAACTCCTAGCCGCTTGTTTTGCTCGCAGCMGGTCTGGGGCRAAACTCATCG
  14  ACTGCGGAACTCCTAGCCGCTTGTTTTGCTCGCAGCAGGTCTGGAGCGAACCTCATCG


 Nt1328                                                  1373
BHCC  GGACTGACAATTCTGTCGTGCTCTCCCGCAAGTATACATCRTTTCC
  14  GGACTGACAATTCTGTCGTGCTCTCCCGCAAGTATACATCATTCCC
```

Alignment of sequences from chronic carriers with the B control consensus sequence in carriers (B). The sequence from patient number 727 is most similar to the consensus sequence. It was chosen as the template for preS1 promoter (nt. 2706-2887).

```
Nt2706                                                          2763
  B  CCAGAGYATGTAGTTAATCATTACTTCCAGACGMGACATTATTTACACACTCTTTGGA
727  CCAGAGTATGTAGTTAATCATTACTTCCAGACGCGACATTATTTACACACTCTTTGGA

Nt2764                                                          2821
  B  AGGCGGGKATCTTATATAAAAGAGAGTCCACACGTAGCGCCTCATTTTGCGGGTCACC
 727 AGGCGGGGATCTTATATAAAAGAGAGTCCACACGTAGCGCCTCATTTTGCGGGTCACC

Nt2822                                                          2879
  B  ATATTCTTGGGAACAAGATCTACAGCATGGGAGGTTGGTCTTCCAAACCTCGAAAAGG
727  ATATTCTTGGGAACAAGATCTACAGCATGGGAGGTTGGTCTTCCAAACCTCGAAAAGG

Nt2880      2887
  B  CATGGGGA
727  CATGGGGA
```

Alignment of sequences from HCC patients with the HBV genotype B consensus sequence in HCC (BHCC). The sequence from patient number 13 is most similar to the consensus sequence (H2). It differs with that in patient number 727 in carrier by only 1 nucleotide, which is the site of mutation at nt. 2712.

```
 Nt2706                                                              2763
BHCC  CCAGAGYATGTAGTTAATCATTACTTCCAGACGMGACATTATTTACACACTCTTTGGA
  13  CCAGAGCATGTAGTTAATCATTACTTCCAGACGCGACATTATTTACACACTCTTTGGA


 Nt2764                                                              2821
BHCC  AGGCGGGRATCTTATATAAAAGAGAGTCCACACGTAGCGCCTCATTTTGCGGGTCACC
  13  AGGCGGGGATCTTATATAAAAGAGAGTCCACACGTAGCGCCTCATTTTGCGGGTCACC


 Nt2822                                                              2879
BHCC  ATATTCTTGGGAACAAGATCTACAGCATGGGAGGTTGGTCTTCCAAACCTCGAAAAGG
  13  ATATTCTTGGGAACAAGATCTACAGCATGGGAGGTTGGTCTTCCAAACCTCGAAAAGG


 Nt2880      2887
BHCC  CATGGGGA
  13  CATGGGGA
```

Alignment of sequences from chronic carriers with the HBV subgenotype Cs consensus sequence in carriers (Cs). The sequence from patient number 1644 is most similar to the consensus sequence. It was chosen as the template for core promoter (nt. 1575-1813).

```
Nt1575                                                                1632
  Cs  CCGTGTGCACTTCGCTTCACCTCTGCACGTCGCATGGAGACCACCGTGAACGCCCGMC
1644  CCGTGTGCACTTCGCTTCACCTCTGCACGTCGCATGGAGACCACCGTGAACGCCCGCC

Nt1633                                                                1690
  Cs  AGGTCTTGCCYAAGGTCTTACATAAGAGGACTCTTGGACTCTCAGCAATGTCAACGAC
1644  AGGTCTTGCCCAAGGTCTTACATAAGAGGACTCTTGGACTCTCAGCAATGTCAACGAC

Nt1691                                                        1748
  Cs  CGACCTTGAGGCATACTTCAAAGACTGTGTATTTAARGACTGGGAGGAGTTGGGGGAG
1644  CGACCTTGAGGCATACTTCAAAGACTGTGTATTTAAAGACTGGGAGGAGTTGGGGGAG

Nt1749                                                                1806
  Cs  GAGAYTAGRTTAAWGATCTTTGTACTGGGAGGCTGTAGGCATAAATTGGTCTGTTCAC
1644  GAGACTAGATTAATGATCTTTGTACTGGGAGGCTGTAGGCATAAATTGGTCTGTTCAC

Nt1807     1813
  Cs  CAGCACC
1644  CAGCACC
```

Alignment of sequences from HCC patient with the HBV subgenotype Cs consensus sequence in HCC (CsHCC). The sequence from patient number 541 is most similar to the consensus sequence. It was chosen as the template for core promoter (nt. 1575-1813).

```
  Nt1575                                                          1632
CsHCC    CCGTGTGCACTTCGCTTCACCTCTGCACGTCGCATGGARACCACCGTGAACGCCCGCC
  541  CCGTGTGCACTTCGCTTCACCTCTGCACGTCGCATGGAAACCACCGTGAACGCCCGCC


  Nt1632                                                          1690
CsHCC  AGGTCTTGCCYAAGGTCTTACATAAGAGGACTCTTGGACTCTCAGCAATGTCAACGAC
  541  AGGTCTTGCCCAAGGTCTTACATAAGAGGACTCTTGGACTCTCAGCAATGTCAACGAC


  Nt1691                                                          1748
CsHCC  CGACCTTGAGGCATACTTCAAAGACTGTGTATTTAARGACTGGGAGGAGTTGGGGGAG
  541  CGACCTTGAGGCATACTTCAAAGACTGTGTATTTAAAGACTGGGAGGAGTTGGGGGAG


   Nt1749                                                         1806
CsHCC  GAGAYTAGRTTAATGATCTTTGTACTRGGAGGCTGTAGGCATAAATTGGTCTGTTCAC
  541  GAGATTAGGTTAATGATCTTTGTACTAGGAGGCTGTAGGCATAAATTGGTCTGTTCAC


  Nt1807      1813
CsHCC  CAGCACC
  541  CAGCACC
```

Alignment of core promoter (nt. 1575-1813) sequences from chronic carriers with the HBV genotype B consensus sequence in carriers (B). The sequence from patient number 1390 is most similar to the consensus sequence. It was chosen as the template for core promoter.

```
   Nt1575            1588
B     CCGTGT GCACTTCG
1390  CCGTGT GCACTTCG


 Nt1589                                                        1638
B     CT TCACCTCTGC ACGTCGCATG GAGACCACCG TGAACGCCCA CGGGAACC
1390  CT TCACCTCTGC ACGTCGCATG GAGACCACCG TGAACGCCCA CGGGAACC


 Nt1639                                                         1688
B     TG CCCAAGGTCT TGCATAAGAG GACTCTTGGA CTTTCAGCAA TGTCAACG
1390  TG CCCAAGGTCT TGCATAAGAG GACTCTTGGA CTTTCAGCAA TGTCAACG


 Nt1689                                                        1738
B     AC CGACCTTGAG GCATACTTCA AAGACTGTGT GTTTAATGAG TGGGAGGA
1390  AC CGACCTTGAG GCATACTTCA AAGACTGTGT GTTTACTGAG TGGGAGGA


 Nt1739                                                        1788
B     GT TGGGGGAGGA GGTTAGGTTA AAGGTCTTTG TACTAGGAGG CTGTAGGC
1390  GT TGGGGGAGGA GATTAGGTTA AAGGTCTTTG TACTAGGAGG CTGTAGGC


 Nt1789                          1813
B     AT AAATTGGTGT GTTCACCAGC ACC
1390  AT AAATTGGTGT GTTCACCAGC ACC
```

Alignment of core promoter (nt. 1575-1813) sequences from chronic carriers with the HBV subgenotype Ce consensus sequence in carriers (Ce). The sequence from patient number 343 is most similar to the consensus sequence. It was chosen as the template for core promoter.

```
 Nt1575                                                              1624
Ce    CCGTGT GCACTTCGCT TCACCTCTGC ACGTCGCATG GARACCACCG TGAA
343   CCGTGT GCACTTCGCT TCACCTCTGC ACGTCGCATG GAGACCACCG TGAA

 Nt1625                                                              1674
Ce    CGCCCA CCAGGTCTTG CCCAAGGTCT TAYATAAGAG GACTCTTGGA CTCY
343   CGCCCA CCAGGTCTTG CCCAAGGTCT TACATAAGAG GACTCTTGGA CTCT

 Nt1675                                                              1724
Ce    CAGCAA TGTCAACGAC CGACCTTGAG GCATACTTCA AAGACTGTKT GTTT
343   CAGCAA TGTCAACGAC CGACCTTGAG GCATACTTCA AAGACTGTGT GTTT

 Nt1725                                                              1774
Ce    AARGAC TGGGAGGAGT TGGGGGAGGA GAYTAGGTTA AWGATCTTTG TACT
343   AAAGAC TGGGAGGAGT TGGGGGAGGA GATTAGGTTA AAGGTCTTTG TACT

 Nt1775                                                 1813
Ce    AGGAGG CTGTAGGCAT AAATTGGTCT GTTCACCAGC ACC
343   AGGAGG CTGTAGGCAT AAATTGGTCT GTTCACCAGC ACC
```

Alignment of 803 HBV NRE sequences (HBV genotype C) from nucleotide 1604 to 1636 (including the RFX1 binding site (nt.1605-1617) and NREBP binding site (nt. 1611-1619). The nt. 1613 is underlined. The HBV sequences were obtained from National Center for Biotechnology Information (GenBank) database (www ncbi nlm nih gov).

```
                                Nt1604                                1636
gi|221228736|gb|FJ518810.1|Hep  TCGCATGGAGACCACCGTGAACGCCCTCCAGGT
gi|221228752|gb|FJ518813.1|Hep  TCGCATGGAGACCACCGTGAACGCCCATCAGGT
gi|260184344|gb|GQ377632.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|94468018|gb|DQ478900.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|94467943|gb|DQ478885.1|Hepa  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|94468013|gb|DQ478899.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|253560603|gb|GQ227695.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|253560589|gb|GQ227693.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|253560609|gb|GQ227696.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|253560596|gb|GQ227694.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|253971057|gb|GQ259588.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|253560616|gb|GQ227697.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|253560583|gb|GQ227692.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|94468008|gb|DQ478898.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|94467933|gb|DQ478883.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|94467978|gb|DQ478892.1|Hepa  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|94467953|gb|DQ478887.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|94467963|gb|DQ478889.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|94468003|gb|DQ478897.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|94467958|gb|DQ478888.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|94467988|gb|DQ478894.1|Hepa  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|94467983|gb|DQ478893.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|94467993|gb|DQ478895.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|94467973|gb|DQ478891.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|94467998|gb|DQ478896.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|94467938|gb|DQ478884.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|94467923|gb|DQ478881.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCTGGT
gi|94467928|gb|DQ478882.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|94467948|gb|DQ478886.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|256029024|gb|GQ372968.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|260184149|gb|GQ377591.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|260183860|gb|GQ377529.1|Hep  TCGCATGGAGACCACCGTGAACGTCCACCAGGT
gi|260183999|gb|GQ377560.1|Hep  TCGCATGGAGACCACCGTGAACAACCGCCAGGT
gi|189171097|gb|EU498227.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|281191015|gb|GQ924643.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|164452829|gb|EU305542.1|Hep  TCGCATGGAAACCACCGTGAACGCCCGACAGGT
gi|164452821|gb|EU305541.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGACAGGT
gi|164452813|gb|EU305540.1|Hep  TCGCACGGAGACCACCGTGAACGCCCGACAGGT
gi|94468022|gb|DQ478901.1|Hepa  TCGCATGGAGACCACCGTGAACGCTCGACAGGT
gi|227335919|gb|FJ904423.1|Hep  TCGCATGGAGACCACCGTGAACGTCCGCCAGGT
gi|281191079|gb|GQ924655.1|Hep  TCGCATGGAAACCACCGTGAACGCCCGCCAGGT
gi|254732794|gb|GQ184326.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|254732790|gb|GQ184325.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|86515926|gb|DQ361534.1|Hepa  TCGCATGGAAACCACCGTGAACGCCCGCCAGGT
gi|281191051|gb|GQ924650.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|281191045|gb|GQ924649.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|281191095|gb|GQ924658.1|Hep  TCGCATGGAAACCACCGTGAACGCCTGCCAGGT
gi|94467968|gb|DQ478890.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCTGGT
gi|207559618|gb|EU833891.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCTGGT
gi|281191090|gb|GQ924657.1|Hep  TCGCATGGAAACCACCGTGAGCGCCCACCAGGC
gi|70610057|gb|DQ089789.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
```

```
                                  Nt1604                                1636
gi|70610052|gb|DQ089788.1|Hepa  TCGCATGGAGACCACCGTGAACGCCTGCCAGGT
gi|70610047|gb|DQ089787.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|70610062|gb|DQ089790.1|Hepa  TCGCATGGARACCACCGTGAACGCCCGCCAGGT
gi|70610042|gb|DQ089786.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|70610037|gb|DQ089785.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|70610027|gb|DQ089783.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|70610022|gb|DQ089782.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|70610037|gb|DQ089785.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|70610017|gb|DQ089781.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|70610007|gb|DQ089779.1|Hepa  TCGCATGGARACCACCGTGAACGCCCGCCAGGT
gi|70610012|gb|DQ089780.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|70610002|gb|DQ089778.1|Hepa  TCGCATGGAAACCACCGTGAACGCCCGCCAGGT
gi|70609997|gb|DQ089777.1|Hepa  TCGCCTGGAGACCACCGTGAACGCCCGCCAGGT
gi|70609992|gb|DQ089776.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|70609987|gb|DQ089775.1|Hepa  TCGCATGGAAACCACCGTGAACGCCCGCCAGGC
gi|70609982|gb|DQ089774.1|Hepa  TCGCATGGAGACCACCGTGAACGCTCGACAGGT
gi|70609977|gb|DQ089773.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|70609972|gb|DQ089772.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|70609962|gb|DQ089770.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGACAGGT
gi|70609957|gb|DQ089769.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGACAGGT
gi|70609938|gb|DQ089765.1|Hepa  TCGCATGGAGACCACCGTGAACACCCGACAGGT
gi|70609933|gb|DQ089764.1|Hepa  TCGCATGGAGACCACCGTGAACGCCAGACAGGT
gi|70609928|gb|DQ089763.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCAACAGGT
gi|70609908|gb|DQ089759.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGACAGGT
gi|70609913|gb|DQ089760.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGACAGGT
gi|70609948|gb|DQ089767.1|Hepa  TCGCATGGAGACCACCGTGAACACCCGACAGGT
gi|70609923|gb|DQ089762.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|70609918|gb|DQ089761.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGACAGGC
gi|70609943|gb|DQ089766.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|70609893|gb|DQ089756.1|Hepa  TCGCATGGAAACCACCGTGAACGCCCGCCAGGC
gi|32346264|gb|AY167092.1|Hepa  TCGCATGGAAACCACCGTGAACGCCCGCCAGGT
gi|70609967|gb|DQ089771.1|Hepa  TCGCATGGAAACCACCGTGAACGTCCGCCAGGT
gi|70609903|gb|DQ089758.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGCCAAGT
gi|70609898|gb|DQ089757.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGACAGGT
gi|70609952|gb|DQ089768.1|Hepa  TTGCATGGAGACCACCGTGAACGCCCGACAAGT
gi|32346292|gb|AY167099.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|269994036|dbj|AB471850.1|He  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|269994026|dbj|AB471848.1|He  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|269994031|dbj|AB471849.1|He  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|32250975|gb|AY306136.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|32346257|gb|AY167090.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|30026448|gb|AY206392.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|30026420|gb|AY206388.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGCCAAGT
gi|32346280|gb|AY167096.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|32346276|gb|AY167095.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|30026426|gb|AY206389.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGC
gi|60460475|gb|AY800249.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|55420294|gb|AY641563.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|55420285|gb|AY641561.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|269994051|dbj|AB471853.1|He  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|269994046|dbj|AB471852.1|He  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|269994041|dbj|AB471851.1|He  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|55420271|gb|AY641558.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|55420290|gb|AY641562.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|55420280|gb|AY641560.1|Hepa  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|55420276|gb|AY641559.1|Hepa  TCGCATGGAAACCACCGTGAACGCCCATCCGGT
gi|32346260|gb|AY167091.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCATCACGT
gi|30026455|gb|AY206393.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
```

```
                                Nt1604                                  1636
gi|30026406|gb|AY206386.1|Hepa  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|30026398|gb|AY206385.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|30026390|gb|AY206384.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|18252533|gb|AF458664.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|15778332|gb|AF411411.1|Hepa  TCGCATGGAAACCACCGTGAACGCCCACTAGGT
gi|15778318|gb|AF411408.1|Hepa  TCGCATGGAAACCACCGTGAACGCCCACTAGGT
gi|18032031|gb|AY066028.1|Hepa  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|15778322|gb|AF411409.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|15778327|gb|AF411410.1|Hepa  TCGCATGGAAACCACCGTGAACGCCCACCTGGT
gi|30026368|gb|AY206381.1|Hepa  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|18252538|gb|AF458665.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|23194246|gb|AF533983.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|15778336|gb|AF411412.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGC
gi|30026375|gb|AY206382.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|30026348|gb|AY206378.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|15072539|gb|AY040627.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|13491148|gb|AF330110.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|29123231|gb|AY217373.1|Hepa  TCGCCTGGAAACCACCGTGAACGCCCACCAGGT
gi|30026355|gb|AY206379.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|30026334|gb|AY206376.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|30026320|gb|AY206374.1|Hepa  TCGCATGGAGACCACCGTGAACACCCGCCAGGT
gi|251822234|dbj|AP011107.1|He  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|251822224|dbj|AP011105.1|He  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|251822229|dbj|AP011106.1|He  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|251822239|dbj|AP011108.1|He  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|226877091|gb|FJ349241.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCTGGT
gi|29123241|gb|AY217376.1|Hepa  TCGCATGGAAACCACCGTGAACGCCCAATAGGT
gi|29123225|gb|AY217371.1|Hepa  TCGCACGGAGACCACCGTGAACGCCCAC-AGGT
gi|29123246|gb|AY217378.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|29123236|gb|AY217375.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGACAGGT
gi|12246993|gb|AF223959.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|12246969|gb|AF223956.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|12246961|gb|AF223955.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|29123229|gb|AY217372.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|12246953|gb|AF223954.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|12247001|gb|AF223960.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGACAGGT
gi|12246985|gb|AF223958.1|Hepa  TCGCATGGAGACCACCGTGAACGCTCGACAGGT
gi|12247009|gb|AF223961.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGACAGGT
gi|12246977|gb|AF223957.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|226876974|gb|FJ349225.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|225421334|gb|FJ361772.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|29124875|gb|AY217358.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACGGGAA
gi|29124870|gb|AY217357.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACGGGAA
gi|29124922|gb|AY217369.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACGGGAA
gi|29124907|gb|AY217365.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACGGGAA
gi|29124886|gb|AY217360.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCGGAA
gi|29124880|gb|AY217359.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCGGAA
gi|29124913|gb|AY217367.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACAGGAA
gi|29124911|gb|AY217366.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACAGGAA
gi|29124916|gb|AY217368.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACAGGAA
gi|29124865|gb|AY217356.1|Hepa  TCGCATGGAAACCACCGTGAACGCCCACAGGAA
gi|29124860|gb|AY217355.1|Hepa  TCGCATGGAAACCACCGTGAACGCCCACAGGAA
gi|29124896|gb|AY217362.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCGGAA
gi|29124892|gb|AY217361.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCGGAA
gi|29124925|gb|AY217370.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACGGGAA
```

```
                                Nt1604                                  1636
gi|29124902|gb|AY217364.1|Hepa  TCGCATGGAAACCACCGTGAACGCCCACGGGAA
gi|18252568|gb|AF461362.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACAGGAA
gi|251822189|dbj|AP011098.1|He  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|195957309|gb|EU916216.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|195957478|gb|EU916241.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|195957462|gb|EU916239.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|195957470|gb|EU916240.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|149211614|gb|DQ993692.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|149211607|gb|DQ993691.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|149211621|gb|DQ993693.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|195957293|gb|EU916213.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|195957282|gb|EU916211.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|195957297|gb|EU916214.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|195957287|gb|EU916212.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|195957454|gb|EU916238.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|195957446|gb|EU916237.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|195957438|gb|EU916236.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|195957344|gb|EU916222.1|Hep  TCGCATGGAGGCCACCGTGAACGCCCACCAGGT
gi|195957331|gb|EU916219.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|195957340|gb|EU916221.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|195957339|gb|EU916220.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|195957352|gb|EU916223.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|195957368|gb|EU916225.1|Hep  TCGCACGGAGACCACCGTGAACGCCCGCCAGGT
gi|195957411|gb|EU916232.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|195957376|gb|EU916226.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|195957360|gb|EU916224.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGAT
gi|149211598|gb|DQ993689.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|149211593|gb|DQ993688.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|195957274|gb|EU916210.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|195957266|gb|EU916207.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|195957269|gb|EU916208.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|195957271|gb|EU916209.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|195957402|gb|EU916230.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|195957434|gb|EU916235.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|195957419|gb|EU916233.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|195957427|gb|EU916234.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|195957260|gb|EU916205.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|195957257|gb|EU916204.1|Hep  TCGCGTGGAGACCACCGTGAACGCCCACCAGGT
gi|195957262|gb|EU916206.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|222142375|gb|FJ562340.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|195957407|gb|EU916231.1|Hep  TCGCATGGAAACCACCATGAACGCCCACCAGGT
gi|195957316|gb|EU916217.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|195957301|gb|EU916215.1|Hep  TCGCATGGAGACCACCGTGAGCGCCCACCAGGT
gi|195957324|gb|EU916218.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|195957382|gb|EU916227.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|195957397|gb|EU916229.1|Hep  TCGCATGGAGACCACCGTGAACGCCCATCAGGT
gi|195957389|gb|EU916228.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|149211604|gb|DQ993690.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|251822184|dbj|AP011097.1|He  TTGCATGGAGACCACCGTGAACGCCCGACAGGT
gi|251822219|dbj|AP011104.1|He  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|251822214|dbj|AP011103.1|He  TCGCATGGAGACCACCGTGAACGCCCACATGGT
gi|251822209|dbj|AP011102.1|He  TCGCATGGAGACCACCGTGAACGCCCACTTGGT
gi|251822204|dbj|AP011101.1|He  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|251822194|dbj|AP011099.1|He  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|251822199|dbj|AP011100.1|He  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
```

```
                                Nt1604                                  1636
gi|222142357|gb|FJ562336.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222142298|gb|FJ562324.1|Hep  TCGCATGGARACCACCGTGAACGCCCACCAGGT
gi|222142313|gb|FJ562327.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|222142352|gb|FJ562335.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222142362|gb|FJ562337.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGY
gi|222142372|gb|FJ562339.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|222142323|gb|FJ562329.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222142333|gb|FJ562331.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222142342|gb|FJ562333.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAAGT
gi|222142328|gb|FJ562330.1|Hep  TCGCATGGARACCACCGTGAACGCCCACCAGGT
gi|222142347|gb|FJ562334.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|222142268|gb|FJ562318.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|222142337|gb|FJ562332.1|Hep  TCGCATGGARACCACCGTGAACGCCCACCWGGT
gi|222142303|gb|FJ562325.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|222142273|gb|FJ562319.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222142278|gb|FJ562320.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222142293|gb|FJ562323.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222142308|gb|FJ562326.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|222142263|gb|FJ562317.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAAGT
gi|222142318|gb|FJ562328.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGC
gi|222142246|gb|FJ562313.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGY
gi|222142224|gb|FJ562308.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222142255|gb|FJ562315.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222142209|gb|FJ562305.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222142219|gb|FJ562307.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222142214|gb|FJ562306.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222142250|gb|FJ562314.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|222142194|gb|FJ562302.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|222142204|gb|FJ562304.1|Hep  TCGCATGGCGACCACCGTGAACGCCCACCAGGT
gi|222142232|gb|FJ562310.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|222142189|gb|FJ562301.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222142179|gb|FJ562299.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|222142153|gb|FJ562293.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222142149|gb|FJ562292.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222142144|gb|FJ562291.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222142158|gb|FJ562294.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|222142171|gb|FJ562297.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222142163|gb|FJ562295.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222142174|gb|FJ562298.1|Hep  TCGCATGGCGACCACCGTGAACGCCCACCAGGT
gi|222142184|gb|FJ562300.1|Hep  TCGCATGGAAACCACCGTGAACGCCCGCCAGGT
gi|222142141|gb|FJ562290.1|Hep  TCGCATGGARACCACCGTGAACGCCCACCAGGT
gi|222142126|gb|FJ562287.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|222142119|gb|FJ562285.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222142114|gb|FJ562284.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222142131|gb|FJ562288.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222142124|gb|FJ562286.1|Hep  TCGCATGGAGACCACCGTGAACGCCCATCAGGT
gi|222142109|gb|FJ562283.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222142104|gb|FJ562282.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222142104|gb|FJ562282.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222142094|gb|FJ562280.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|222142025|gb|FJ562266.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222141918|gb|FJ562244.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222142089|gb|FJ562279.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222141957|gb|FJ562252.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222142065|gb|FJ562274.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
```

```
                                Nt1604                                  1636
gi|222141908|gb|FJ562242.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|222142070|gb|FJ562275.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222142045|gb|FJ562270.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAAGT
gi|222141987|gb|FJ562258.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|222141942|gb|FJ562249.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|222142080|gb|FJ562277.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222142040|gb|FJ562269.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222141952|gb|FJ562251.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222142050|gb|FJ562271.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|222142002|gb|FJ562261.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222142020|gb|FJ562265.1|Hep  TCGCATGGARACCACCGTGAACGCCCACCAGGT
gi|222142055|gb|FJ562272.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222141937|gb|FJ562248.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|222142060|gb|FJ562273.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222142015|gb|FJ562264.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222141913|gb|FJ562243.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|222142035|gb|FJ562268.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|222142075|gb|FJ562276.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222142084|gb|FJ562278.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|222141972|gb|FJ562255.1|Hep  TCGCATGGAGACCACCGTGAACGTCCGCCAGGT
gi|222142030|gb|FJ562267.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|222141923|gb|FJ562245.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|222141977|gb|FJ562256.1|Hep  TCGCATGGAAACCACCGTGAACACCCGCCAGGT
gi|222141947|gb|FJ562250.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222141932|gb|FJ562247.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213495152|gb|FJ386657.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|213495068|gb|FJ386645.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|222141893|gb|FJ562239.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|213495244|gb|FJ386670.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213495197|gb|FJ386663.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|213495123|gb|FJ386653.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222141873|gb|FJ562235.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213495167|gb|FJ386659.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|213494974|gb|FJ386632.1|Hep  TCGCATGGARACCACCGTGAACGCCCACCAGGC
gi|213495189|gb|FJ386662.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213495182|gb|FJ386661.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213495224|gb|FJ386667.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213495008|gb|FJ386637.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213494982|gb|FJ386633.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213494884|gb|FJ386619.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222141888|gb|FJ562238.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|222141835|gb|FJ562227.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|213495104|gb|FJ386650.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|213495096|gb|FJ386649.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|213495266|gb|FJ386673.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222141903|gb|FJ562241.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|222141863|gb|FJ562233.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|213495036|gb|FJ386641.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213495365|gb|FJ386687.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|222141791|gb|FJ562218.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|213495020|gb|FJ386639.1|Hep  TCGCATGGAGACCACCGTGAACGCTCACCAGGT
gi|213495360|gb|FJ386686.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213495352|gb|FJ386685.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222141830|gb|FJ562226.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213494936|gb|FJ386627.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACMAGGT
```

```
                                Nt1604                                  1636
gi|213495310|gb|FJ386679.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213494994|gb|FJ386635.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213495015|gb|FJ386638.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213494959|gb|FJ386630.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213494951|gb|FJ386629.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222141849|gb|FJ562230.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222141858|gb|FJ562232.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|213495210|gb|FJ386665.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213495302|gb|FJ386678.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213495251|gb|FJ386671.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|213494915|gb|FJ386624.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|213495115|gb|FJ386652.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213494966|gb|FJ386631.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|213495052|gb|FJ386643.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|222141825|gb|FJ562225.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213495258|gb|FJ386672.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213495379|gb|FJ386689.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213494892|gb|FJ386620.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|88595348|dbj|AB205152.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|213495080|gb|FJ386647.1|Hep  TCGCATGGAAACCACCGTGAACGCCCTCCAGGT
gi|157679138|dbj|AB298721.1|He  TCGCATGGAGGCCACCGTGAACGCCCACCAGGT
gi|157679133|dbj|AB298720.1|He  TCGCATGGAGGCCACCGTGAACGCCCACCAGGT
gi|222141806|gb|FJ562221.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213494922|gb|FJ386625.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|213495295|gb|FJ386677.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213495060|gb|FJ386644.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222141840|gb|FJ562228.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|222141801|gb|FJ562220.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|213495108|gb|FJ386651.1|Hep  TCGCATGGARACCACCGTGAACGCCCACCAGGT
gi|213494944|gb|FJ386628.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213495202|gb|FJ386664.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213494908|gb|FJ386623.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213494903|gb|FJ386622.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|213494929|gb|FJ386626.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213494900|gb|FJ386621.1|Hep  TCGCATGGCAACCACCGTGAACGCCCACCAGGT
gi|213495028|gb|FJ386640.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCGGAT
gi|213495075|gb|FJ386646.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCGGRA
gi|222141845|gb|FJ562229.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213495273|gb|FJ386674.1|Hep  TCGCATGGAAACCACCGTGATCGCCCACCGGAA
gi|222141815|gb|FJ562223.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213494718|gb|FJ386595.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213494561|gb|FJ386574.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213494597|gb|FJ386579.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292820|gb|EU939642.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|202072691|gb|FJ032360.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|202072684|gb|FJ032359.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|202072695|gb|FJ032361.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292705|gb|EU939614.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292701|gb|EU939613.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|213494876|gb|FJ386618.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213494781|gb|FJ386604.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292876|gb|EU939656.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292872|gb|EU939655.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213494817|gb|FJ386609.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213494611|gb|FJ386581.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
```

```
                                Nt1604                                1636
gi|213494643|gb|FJ386585.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|213494592|gb|FJ386578.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|213494703|gb|FJ386593.1|Hep  TCGCATGGAAACCACCATGAACGCCCACCAGGT
gi|213494697|gb|FJ386592.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292811|gb|EU939640.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|202072552|gb|FJ032341.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|202072544|gb|FJ032340.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|213494746|gb|FJ386599.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGC
gi|197292692|gb|EU939611.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|213494732|gb|FJ386597.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213494738|gb|FJ386598.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292727|gb|EU939619.1|Hep  TCGCATGGAAACCACCATGAACGCCCACCAGGT
gi|213494689|gb|FJ386591.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|202072579|gb|FJ032345.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|213494584|gb|FJ386577.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292832|gb|EU939645.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|202072621|gb|FJ032351.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|202072614|gb|FJ032350.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|202072505|gb|FJ032334.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292687|gb|EU939610.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292722|gb|EU939618.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|202072660|gb|FJ032356.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGC
gi|202072652|gb|FJ032355.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGC
gi|213494840|gb|FJ386612.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|202072529|gb|FJ032338.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|202072516|gb|FJ032336.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|202072523|gb|FJ032337.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|202072536|gb|FJ032339.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGC
gi|197292682|gb|EU939609.1|Hep  TCGCATGGAAACCACCGTGAACGCCCATCAGGT
gi|213494710|gb|FJ386594.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|202072513|gb|FJ032335.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292828|gb|EU939644.1|Hep  TCGCATGGAAACCACCGTGAACGCCCGCCAGGT
gi|197292888|gb|EU939659.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|213494853|gb|FJ386614.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292835|gb|EU939646.1|Hep  TCACATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292880|gb|EU939657.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|202072485|gb|FJ032331.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292697|gb|EU939612.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213494673|gb|FJ386589.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213494846|gb|FJ386613.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213494871|gb|FJ386617.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213494760|gb|FJ386601.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292843|gb|EU939648.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292816|gb|EU939641.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213494866|gb|FJ386616.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213494568|gb|FJ386575.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213494803|gb|FJ386607.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213494767|gb|FJ386602.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292855|gb|EU939651.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292838|gb|EU939647.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213494576|gb|FJ386576.1|Hep  TCGCATGGAGACCTCCGTGAACACCCACCAGGT
gi|197292714|gb|EU939616.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292868|gb|EU939654.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292848|gb|EU939649.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292884|gb|EU939658.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
```

```
                                Nt1604                                  1636
gi|197292859|gb|EU939652.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|213494657|gb|FJ386587.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|202072595|gb|FJ032347.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|202072587|gb|FJ032346.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|202072602|gb|FJ032348.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|202072493|gb|FJ032332.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292852|gb|EU939650.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213494726|gb|FJ386596.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|213494796|gb|FJ386606.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|213494773|gb|FJ386603.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|213494788|gb|FJ386605.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAAGT
gi|197292824|gb|EU939643.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213494665|gb|FJ386588.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292709|gb|EU939615.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|213494833|gb|FJ386611.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292864|gb|EU939653.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|202072499|gb|FJ032333.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292718|gb|EU939617.1|Hep  TCGCATGGAGACCACCATGAACGCCCACCAGGT
gi|125995352|dbj|AB270535.1|He  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|125995350|dbj|AB270534.1|He  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|213494651|gb|FJ386586.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|202072566|gb|FJ032343.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292754|gb|EU939626.1|Hep  TCGCATGGAGACCACCGTGAACGTCCGCCAGGT
gi|197292749|gb|EU939625.1|Hep  TCGCATGGAAACCACCGTGAACGTCCACCAGGT
gi|197292746|gb|EU939624.1|Hep  TCGCATGGAAACCACCGTGAACACCCGCCAGGT
gi|213494605|gb|FJ386580.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292742|gb|EU939623.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292732|gb|EU939620.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292740|gb|EU939622.1|Hep  TCGCATGGCAACCACCGTGAACGCCCACCAGGT
gi|197292736|gb|EU939621.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGC
gi|197292681|gb|EU939608.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGC
gi|197292643|gb|EU939599.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292673|gb|EU939606.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292507|gb|EU939566.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292613|gb|EU939592.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292552|gb|EU939578.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292561|gb|EU939580.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292527|gb|EU939571.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292635|gb|EU939597.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292677|gb|EU939607.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292618|gb|EU939593.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292537|gb|EU939573.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292569|gb|EU939582.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|197292519|gb|EU939569.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|197292532|gb|EU939572.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292594|gb|EU939588.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292582|gb|EU939585.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292574|gb|EU939583.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292652|gb|EU939601.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAAGT
gi|197292590|gb|EU939587.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292557|gb|EU939579.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292549|gb|EU939577.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292656|gb|EU939602.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292578|gb|EU939584.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292668|gb|EU939605.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
```

```
                                Nt1604                                  1636
gi|197292608|gb|EU939591.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292603|gb|EU939590.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292492|gb|EU939562.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292487|gb|EU939561.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCGGGT
gi|197292483|gb|EU939560.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292627|gb|EU939595.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292598|gb|EU939589.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292505|gb|EU939565.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292545|gb|EU939576.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292523|gb|EU939570.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292541|gb|EU939574.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292664|gb|EU939604.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292496|gb|EU939563.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292640|gb|EU939598.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292543|gb|EU939575.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292623|gb|EU939594.1|Hep  TCGCATGGAGACCCCCGTGAACGCCCACCAGGT
gi|197292586|gb|EU939586.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGC
gi|197292511|gb|EU939567.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGC
gi|197292647|gb|EU939600.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292564|gb|EU939581.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292659|gb|EU939603.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGC
gi|197292500|gb|EU939564.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|197292631|gb|EU939596.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292515|gb|EU939568.1|Hep  TCGCATGGARACCACCGTGAACGCCCACCAGGT
gi|197292478|gb|EU939559.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACAGGAA
gi|197292681|gb|EU939608.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGC
gi|197292643|gb|EU939599.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292673|gb|EU939606.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292507|gb|EU939566.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292613|gb|EU939592.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292552|gb|EU939578.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292561|gb|EU939580.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292527|gb|EU939571.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292635|gb|EU939597.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292677|gb|EU939607.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292618|gb|EU939593.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292537|gb|EU939573.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292569|gb|EU939582.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|197292519|gb|EU939569.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|197292532|gb|EU939572.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292594|gb|EU939588.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292582|gb|EU939585.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292574|gb|EU939583.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292652|gb|EU939601.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAAGT
gi|197292590|gb|EU939587.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292557|gb|EU939579.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292549|gb|EU939577.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292656|gb|EU939602.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292578|gb|EU939584.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292668|gb|EU939605.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292608|gb|EU939591.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292603|gb|EU939590.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292492|gb|EU939562.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292487|gb|EU939561.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCGGGT
gi|197292483|gb|EU939560.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
```

```
                                Nt1604                                1636
gi|197292627|gb|EU939595.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292598|gb|EU939589.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292505|gb|EU939565.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292545|gb|EU939576.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292523|gb|EU939570.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292541|gb|EU939574.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292664|gb|EU939604.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292496|gb|EU939563.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292640|gb|EU939598.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292543|gb|EU939575.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292623|gb|EU939594.1|Hep  TCGCATGGAGACCCCCGTGAACGCCCACCAGGT
gi|197292586|gb|EU939586.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGC
gi|197292511|gb|EU939567.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGC
gi|197292647|gb|EU939600.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|197292564|gb|EU939581.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292659|gb|EU939603.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGC
gi|197292500|gb|EU939564.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|197292631|gb|EU939596.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|197292515|gb|EU939568.1|Hep  TCGCATGGARACCACCGTGAACGCCCACCAGGT
gi|197292478|gb|EU939559.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACAGGAA
gi|182260489|gb|EU570074.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|182260463|gb|EU570073.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|182260282|gb|EU570068.1|Hep  TCGCCTGGAGACCACCGTGAACACCCGCCAGGT
gi|182260246|gb|EU570067.1|Hep  TCGCATGGAGACCACCGTGAACACCCGCCTGGT
gi|172052353|gb|EU560440.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|172034785|gb|EU562219.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|172034777|gb|EU562218.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|172052361|gb|EU560441.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|169403920|dbj|AB426467.1|He  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|168495255|gb|EU439016.1|Hep  TCGCATGGAAACCACCGTGAACGCCCATCAGGT
gi|168495250|gb|EU439015.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|172034769|gb|EU562217.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|172034753|gb|EU562215.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|172034761|gb|EU562216.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|164654807|gb|EU306728.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|164654791|gb|EU306726.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|164654778|gb|EU306724.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|171769134|gb|EU439009.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|171769131|gb|EU439014.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCGGGT
gi|164654815|gb|EU306729.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|164654783|gb|EU306725.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|164654772|gb|EU306723.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|168495216|gb|EU439005.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|164654799|gb|EU306727.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|164654764|gb|EU306722.1|Hep  TCGCATGGAGACCACCGTGAACACCCACCAGGT
gi|171769129|gb|EU439013.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|168495226|gb|EU439007.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|164654735|gb|EU306718.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|164654756|gb|EU306721.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|164654748|gb|EU306720.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|164654740|gb|EU306719.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|168495221|gb|EU439006.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|168495294|gb|EU439025.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|168495245|gb|EU439012.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|168495230|gb|EU439008.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
```

```
                                Nt1604                                1636
gi|168495240|gb|EU439011.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|168495235|gb|EU439010.1|Hep  TCGCATGGAGACCACCGTGGACGCCCACCAGGT
gi|182260426|gb|EU570072.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|172052346|gb|EU560439.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|172052338|gb|EU560438.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|171345901|gb|EU547562.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|171345894|gb|EU547561.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|156571791|gb|EF494377.1|Hep  TCGAATGGCGACCACCGTGAACGCCCACCAGGT
gi|171345878|gb|EU547559.1|Hep  TCGCATGGCGACCACCGTGAACGCCCACCAGGT
gi|171345886|gb|EU547560.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGACAAGT
gi|156571806|gb|EF494379.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGACAGGT
gi|171345870|gb|EU547558.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|168495259|gb|EU439017.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCTGGT
gi|156571784|gb|EF494376.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCATGT
gi|156571799|gb|EF494378.1|Hep  TTGCATGGAGACCACCGTGAACGCCCATCAGAT
gi|164654726|gb|EU306716.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|164654721|gb|EU306715.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|164654730|gb|EU306717.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|164654713|gb|EU306714.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|164654705|gb|EU306713.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|164654543|gb|EU306692.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|164654527|gb|EU306690.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|164654535|gb|EU306691.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|164654549|gb|EU306693.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|164654705|gb|EU306713.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|164654519|gb|EU306689.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|164654511|gb|EU306688.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|164654503|gb|EU306687.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|164654489|gb|EU306685.1|Hep  TCGCATGGAGACCACCGTGAACGCCTGCCAGGT
gi|164654497|gb|EU306686.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|151936657|gb|EF688062.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|39748631|dbj|AB112063.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|164654429|gb|EU306675.1|Hep  TTGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|164654423|gb|EU306674.1|Hep  TTGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|164654415|gb|EU306673.1|Hep  TTGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|148524698|dbj|AB195957.1|He  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|148524693|dbj|AB195956.1|He  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|148524688|dbj|AB195955.1|He  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|119672884|dbj|AB288026.1|He  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|155968653|dbj|AB202072.1|He  TCGCATGGAGGCCACCGTGAACGCCCACCAGGT
gi|155968648|dbj|AB202071.1|He  TCGCATGGAGGCCACCGTGAACGCCCACCAGGT
gi|160358291|dbj|AB368296.1|He  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|160358296|dbj|AB368297.1|He  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|164654409|gb|EU306672.1|Hep  TTGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|164654402|gb|EU306671.1|Hep  TTGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|148524683|dbj|AB195954.1|He  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|148524678|dbj|AB195953.1|He  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|148524673|dbj|AB195952.1|He  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|148524668|dbj|AB195951.1|He  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|148524663|dbj|AB195950.1|He  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|148524658|dbj|AB195949.1|He  TAGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|146217462|gb|EF536066.1|Mut  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|146217454|gb|EF536065.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|148524618|dbj|AB195941.1|He  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|148524613|dbj|AB195940.1|He  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
```

```
                                Nt1604                                  1636
gi|148524608|dbj|AB195939.1|He  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|148524653|dbj|AB195948.1|He  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|148524648|dbj|AB195947.1|He  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|148524633|dbj|AB195944.1|He  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|148524628|dbj|AB195943.1|He  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|148524623|dbj|AB195942.1|He  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|84095091|dbj|AB182589.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|110613494|gb|DQ683578.1|Hep  TCGCATGGAAACCACCGTGAACGCCCATCAGGT
gi|110456033|gb|DQ536412.1|Hep  TCGCATGGAGACCACCATGAACGCCCATCTGGT
gi|110456030|gb|DQ536410.1|Hep  TCGCATGGAGACCACCATGAACGCCCATCTGGT
gi|148524643|dbj|AB195946.1|He  ACGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|148524638|dbj|AB195945.1|He  ACGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|148524603|dbj|AB195938.1|He  TCGCATGGAAACCACCGTGAACGCCCGCCAAGT
gi|148524598|dbj|AB195937.1|He  TCGCATGGAAACCACCGTGAACGCCCGCCAAGT
gi|148524593|dbj|AB195936.1|He  TCGCATGGAAACCACCGTGAACGCCCGCCAGGT
gi|87886469|dbj|AB250109.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|148524573|dbj|AB195932.1|He  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|148524568|dbj|AB195931.1|He  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|148524563|dbj|AB195930.1|He  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|87080764|dbj|AB247916.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|74474927|dbj|AB206817.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|74474922|dbj|AB206816.1|Hep  TCGCATGGAAACCACCGTGAACGCCCGCCAGGT
gi|110456038|gb|DQ536414.1|Hep  TCGCATGGAGACCACCATGAACGCCCATCTGGT
gi|110456037|gb|DQ536413.1|Hep  TCGCATGGAGACCACCATGAACGCCCATCTGGT
gi|110456032|gb|DQ536411.1|Hep  TCGCATGGAGACCACCATGAACGCCCATCTGGT
gi|95007765|dbj|AB241113.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|95007760|dbj|AB241112.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|95007755|dbj|AB241111.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|95007750|dbj|AB241110.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|95007745|dbj|AB241109.1|Hep  TTGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|60098337|emb|AJ748098.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|39979297|dbj|AB112472.1|Hep  AGGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|60279615|dbj|AB205125.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|29123245|gb|AY217377.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|39748636|dbj|AB112065.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|42558175|dbj|AB111946.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|39930180|dbj|AB112408.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAAGT
gi|39840922|dbj|AB112348.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|39748641|dbj|AB112066.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|39979292|dbj|AB112471.1|Hep  TCGCATGGAGACCACCGTGAACGCCTGCCAGGT
gi|29123235|gb|AY217374.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|11071965|dbj|AB042283.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|11071963|dbj|AB042282.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|60279610|dbj|AB205124.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|6063437|dbj|AB033553.1|Hepa  TCGCATGGAAACCACCGTGAACACCCACCAGGT
gi|6063452|dbj|AB033556.1|Hepa  TCGCATGGAAACCACCGTGAACACCCACCAGGT
gi|6063432|dbj|AB033552.1|Hepa  TCGCATGGAAACCACCGTGAACACCCACCTGGT
gi|6063427|dbj|AB033551.1|Hepa  TCGCATGGAAACCACCGTGAACACCCACCAGGT
gi|6063422|dbj|AB033550.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|56342131|dbj|AB113879.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|11071969|dbj|AB042285.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|18252573|gb|AF461363.1|AF46  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|18252543|gb|AF461357.1|AF46  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|6009775|dbj|AB026814.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|6009770|dbj|AB026813.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
```

```
                                Nt1604                                  1636
gi|6009765|dbj|AB026812.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|6009760|dbj|AB026811.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|18252553|gb|AF461359.1|AF46  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|18252548|gb|AF461358.1|AF46  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|60279605|dbj|AB205123.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|18252563|gb|AF461361.1|AF46  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|11071967|dbj|AB042284.1|Hep  TCGCATGGCAACCACCGTGAACGCCCACCAGGT
gi|56342116|dbj|AB113876.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|56342111|dbj|AB113875.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|56342121|dbj|AB113877.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|6063457|dbj|AB033557.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|56342126|dbj|AB113878.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|42558181|dbj|AB115417.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|38603641|dbj|AB115418.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|13365550|dbj|AB048705.1|Hep  TCGCATGGAGACCACCGTGAACACCCACATGAT
gi|13365548|dbj|AB048704.1|Hep  TCGCATGGAGACCACCGTGAACACCCACATGAT
gi|21280247|dbj|AB073831.1|Hep  TCGCATGGAGACCACCGTGAACGCTCACCGGAA
gi|21280245|dbj|AB073830.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACTGGAA
gi|29124901|gb|AY217363.1|Hepa  TCGCATGGAAACCACCGTGAACGCCCACTGGAA
gi|21280249|dbj|AB073832.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCGGAA
gi|21280253|dbj|AB073834.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACGGGAA
gi|21280241|dbj|AB073828.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACGGGAA
gi|21280239|dbj|AB073827.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACGGGAA
gi|21280251|dbj|AB073833.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACGGGAA
gi|21280265|dbj|AB073840.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACGGGAA
gi|21280237|dbj|AB073826.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACAGGAA
gi|21280229|dbj|AB073822.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACGGGAA
gi|21280243|dbj|AB073829.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACGGGAA
gi|21280259|dbj|AB073837.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACTGGAA
gi|21280231|dbj|AB073823.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACAGAAA
gi|21280263|dbj|AB073839.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACGGGAA
gi|21280233|dbj|AB073824.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACGGGAA
gi|21280267|dbj|AB073841.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACGGGAA
gi|21280227|dbj|AB073821.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACAKTAA
gi|21280257|dbj|AB073836.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACTGGAA
gi|21280235|dbj|AB073825.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACTGGAA
gi|21280255|dbj|AB073835.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCGGAA
gi|21280299|dbj|AB073857.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCGGAA
gi|21280297|dbj|AB073856.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCGGAA
gi|21280287|dbj|AB073851.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCGGAA
gi|21280277|dbj|AB073846.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCGGAA
gi|21280281|dbj|AB073848.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCGGAA
gi|21280261|dbj|AB073838.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCGGAA
gi|21280273|dbj|AB073844.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCGGAA
gi|21280271|dbj|AB073843.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCGGAA
gi|21280275|dbj|AB073845.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCGGAA
gi|21280269|dbj|AB073842.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCGGAA
gi|21280285|dbj|AB073850.1|Hep  TCGCATGGAAACCACCGTGAACGCCCACCGGAA
gi|21280279|dbj|AB073847.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCGGAA
gi|21280295|dbj|AB073855.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCGGAA
gi|21280293|dbj|AB073854.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCGGAA
gi|21280289|dbj|AB073852.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCGGAA
gi|21280301|dbj|AB073858.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACTGGAA
gi|21280283|dbj|AB073849.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCGGAA
gi|21280291|dbj|AB073853.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCGGAA
```

```
                                Nt1604                                  1636
gi|13365546|dbj|AB048703.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAATT
gi|13365544|dbj|AB048702.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAATT
gi|13365542|dbj|AB048701.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAATT
gi|281190904|gb|GQ924623.1|Hep  TCGCATGGAGACCACCGTGAACGCCTGCCAGGT
gi|281190794|gb|GQ924604.1|Hep  TCGCATGGAGACCACCGTGAACGCCTGCCAGGT
gi|281190876|gb|GQ924618.1|Hep  TCGCATGGAGACCACCGTGAACGCCTGCCAGGT
gi|281190860|gb|GQ924615.1|Hep  TCGCATGGAGACCACCGTGAACGTCTGCCAGGT
gi|281190882|gb|GQ924619.1|Hep  TCGCATGGAGACCACCGTGAACGCCTGCCAGGT
gi|281191009|gb|GQ924642.1|Hep  TCGCATGGAGACCACCGTGAACGCCTGCCAGGT
gi|281190865|gb|GQ924616.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|281190824|gb|GQ924609.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|281190848|gb|GQ924613.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|281190842|gb|GQ924612.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|281190854|gb|GQ924614.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|70610125|gb|DQ089803.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|281190936|gb|GQ924629.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|281190976|gb|GQ924636.1|Hep  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|79013072|gb|DQ246215.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGCCAAGT
gi|70610072|gb|DQ089792.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|70610067|gb|DQ089791.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|70610130|gb|DQ089804.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|70610090|gb|DQ089796.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|70610085|gb|DQ089795.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|281190958|gb|GQ924633.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|70610095|gb|DQ089797.1|Hepa  TCGCATGGAAACCACCGTGAACGCCCACCAGGT
gi|70610082|gb|DQ089794.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|70610077|gb|DQ089793.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|70610100|gb|DQ089798.1|Hepa  TCGCATGGAGACCTCCGTGAACGCCCACCAGGT
gi|70610110|gb|DQ089800.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|70610105|gb|DQ089799.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|70610120|gb|DQ089802.1|Hepa  TCGCATGGAGACCACCGTGAACGCCCGCCAGGT
gi|70610115|gb|DQ089801.1|Hepa  TCGCATGGCGACCACCGTGAACGCCCACCAGGT
gi|281190899|gb|GQ924622.1|Hep  TCGCATGGAGACCACCGTGAACGCCCACCAGGT
gi|281190887|gb|GQ924620.1|Hep  TCGCATGGAGACCACCGTGAACGCCAACCTGGT
```

Dataset of univariate analysis of HBV viral load and G1613A mutation in HBV carriers.

Serum Number	SEX	AFP	Viral load (copies/ml)	HBV genotype	Nucleotide at 1613
1	F	0	479.25	Cs	A
2	F	0	1016.31	Cs	A
3	F	0	1280.36	Cs	G
4	F	0	1776.79	Cs	G
5	F	0	2109.32	Cs	A
6	F	0	2374.64	Cs	G
7	F	0	2742.25	Cs	G
8	F	0	4079.33	Cs	G
9	F	0	7755.13	Cs	G
10	F	0	8151.00	Cs	G
11	F	0	8408.90	Cs	G
12	F	0	10525.72	Cs	G
13	F	0	11307.89	Cs	G
14	F	0	11635.21	Cs	A
15	F	0	13100.00	Cs	G
16	F	0	13858.92	Cs	G
17	F	77	17774.21	Cs	G
18	F	0	22677.87	Cs	G
19	F	0	26373.09	Cs	G
20	F	13	40857.83	Cs	G
21	F	0	57980.60	Cs	G
22	F	0	61375.84	Cs	G
23	F	10	89764.45	Cs	G
24	F	0	116049.87	Cs	G
25	F	17	131614.20	Cs	G
26	F	0	133975.26	Cs	G
27	F	0	148205.28	Cs	G
28	F	20	153666.97	Cs	A
29	F	83	186255.10	Cs	G
30	F	0	203254.56	Cs	G
31	F	0	232478.34	Cs	G
32	F	0	251452.42	Cs	G
33	F	0	286000.00	Cs	G
34	F	0	310015.00	Cs	A
35	F	0	334586.55	Cs	G
36	F	0	361164.50	Cs	G
37	F	0	375644.25	Cs	G
38	F	0	393493.03	Cs	G
39	F	0	483000.00	Cs	G

40	F	0	530000.00	Cs	G
41	F	0	560000.00	Cs	G
42	F	0	590000.00	Cs	G
43	F	0	800000.00	Cs	G
44	F	0	840000.00	Cs	G
45	F	0	915000.00	Cs	G
46	F	10	945000.00	Cs	A
47	F	15	975000.00	Cs	G
48	F	0	1065000.00	Cs	A
49	F	10	1145000.00	Cs	G
50	F	0	1230000.00	Cs	A
51	F	0	1235000.00	Cs	G
52	F	14	1275000.00	Cs	G
53	F	0	1500000.00	Cs	A
54	F	0	1510000.00	Cs	A
55	F	0	1565000.00	Cs	A
56	F	0	1955000.00	Cs	A
57	F	0	2000000.00	Cs	A
58	F	0	2130000.00	Cs	G
59	F	0	2265000.00	Cs	G
60	F	0	2380000.00	Cs	G
61	F	0	2530000.00	Cs	G
62	F	0	2615000.00	Cs	A
63	F	11	2625000.00	Cs	G
64	F	18	3880000.00	Cs	A
65	F	0	3895000.00	Cs	A
66	F	0	4655000.00	Cs	G
67	F	0	4910000.00	Cs	G
68	F	0	5150000.00	Cs	A
69	F	0	5350000.00	Cs	G
70	F	0	5450000.00	Cs	G
71	F	0	6100000.00	Cs	G
72	F	0	6650000.00	Cs	A
73	F	0	7600000.00	Cs	A
74	F	0	11550000.00	Cs	G
75	F	0	12350000.00	Cs	G
76	F	0	12750000.00	Cs	G
77	F	0	18850000.00	Cs	G
78	F	143	19150000.00	Cs	G
79	F	0	27050000.00	Cs	A
80	F	0	41150000.00	Cs	G
81	F	0	46300000.00	Cs	G
82	F	0	122000000.00	Cs	G
83	F	19	130500000.00	Cs	A
84	F	0	165000000.00	Cs	G

85	F	0	181500000.00	Cs	G
86	F	0	259000000.00	Cs	G
87	F	0	391500000.00	Cs	G
88	F	0	580000000.00	Cs	A
89	M	0	163.38	Cs	G
90	M	0	180.75	Cs	G
91	M	0	206.87	Cs	G
92	M	0	430.89	Cs	G
93	M	0	756.90	Cs	G
94	M	0	1057.96	Cs	G
95	M	0	1075.09	Cs	A
96	M	0	1103.60	Cs	G
97	M	0	1271.21	Cs	G
98	M	0	1494.32	Cs	A
99	M	0	1578.01	Cs	G
100	M	0	1807.61	Cs	G
101	M	12	2634.96	Cs	G
102	M	11	2705.89	Cs	G
103	M	0	5034.71	Cs	A
104	M	0	5062.98	Cs	G
105	M	11	8151.63	Cs	G
106	M	0	10288.35	Cs	G
107	M	0	12249.43	Cs	G
108	M	0	12555.93	Cs	G
109	M	0	14171.31	Cs	G
110	M	0	14375.88	Cs	G
111	M	0	14427.36	Cs	G
112	M	10	14461.90	Cs	G
113	M	0	15653.36	Cs	G
114	M	0	15685.07	Cs	G
115	M	0	16862.01	Cs	G
116	M	0	19517.66	Cs	A
117	M	0	20490.69	Cs	G
118	M	16	20510.40	Cs	A
119	M	0	20759.88	Cs	G
120	M	0	21976.71	Cs	G
121	M	0	25442.23	Cs	G
122	M	24	30153.22	Cs	A
123	M	0	35048.80	Cs	G
124	M	0	35139.98	Cs	G
125	M	0	40542.88	Cs	G
126	M	0	64131.45	Cs	G
127	M	0	70858.72	Cs	A
128	M	10	85887.42	Cs	G
129	M	0	87560.83	Cs	G

130	M	0	120782.50	Cs	G
131	M	0	124904.21	Cs	A
132	M	0	127181.03	Cs	A
133	M	0	130000.00	Cs	G
134	M	0	131331.03	Cs	A
135	M	0	134288.16	Cs	G
136	M	0	157378.99	Cs	G
137	M	0	158457.66	Cs	G
138	M	0	162587.50	Cs	G
139	M	0	169500.00	Cs	A
140	M	0	175711.14	Cs	A
141	M	0	178325.00	Cs	G
142	M	0	189376.28	Cs	G
143	M	0	277640.91	Cs	G
144	M	0	283933.86	Cs	G
145	M	0	286405.00	Cs	G
146	M	0	295416.11	Cs	G
147	M	0	314003.55	Cs	G
148	M	0	315943.03	Cs	G
149	M	0	324797.07	Cs	G
150	M	0	331317.33	Cs	A
151	M	0	362372.32	Cs	G
152	M	0	396532.69	Cs	G
153	M	0	401674.57	Cs	G
154	M	0	413745.52	Cs	G
155	M	0	422522.99	Cs	G
156	M	0	444960.49	Cs	A
157	M	13	483915.78	Cs	A
158	M	47	505000.00	Cs	G
159	M	0	590000.00	Cs	G
160	M	0	620000.00	Cs	G
161	M	10	705000.00	Cs	G
162	M	0	725000.00	Cs	G
163	M	0	730000.00	Cs	A
164	M	0	770000.00	Cs	G
165	M	0	805000.00	Cs	G
166	M	19	860000.00	Cs	A
167	M	0	860000.00	Cs	G
168	M	0	1045000.00	Cs	G
169	M	0	1055000.00	Cs	G
170	M	0	1060000.00	Cs	G
171	M	0	1115000.00	Cs	G
172	M	0	1115000.00	Cs	G
173	M	0	1125000.00	Cs	A
174	M	0	1145000.00	Cs	G

175	M	0	1160000.00	Cs	G
176	M	0	1185000.00	Cs	G
177	M	0	1190000.00	Cs	G
178	M	0	1245000.00	Cs	G
179	M	0	1265000.00	Cs	A
180	M	25	1345000.00	Cs	G
181	M	0	1370000.00	Cs	A
182	M	0	1505000.00	Cs	G
183	M	0	1515000.00	Cs	G
184	M	0	1525000.00	Cs	G
185	M	17	1540000.00	Cs	G
186	M	0	1620000.00	Cs	G
187	M	15	1680000.00	Cs	A
188	M	0	1780000.00	Cs	G
189	M	0	1785000.00	Cs	G
190	M	12	1970000.00	Cs	G
191	M	0	2060000.00	Cs	G
192	M	0	2115000.00	Cs	G
193	M	0	2340000.00	Cs	G
194	M	14	2400000.00	Cs	G
195	M	0	2425000.00	Cs	G
196	M	0	2440000.00	Cs	G
197	M	0	2655000.00	Cs	G
198	M	0	2695000.00	Cs	G
199	M	0	2900000.00	Cs	G
200	M	0	3020000.00	Cs	A
201	M	0	3215000.00	Cs	G
202	M	0	3295000.00	Cs	G
203	M	0	3320000.00	Cs	G
204	M	0	3355000.00	Cs	G
205	M	0	3565000.00	Cs	G
206	M	16	3605000.00	Cs	G
207	M	0	3775000.00	Cs	G
208	M	0	4090000.00	Cs	G
209	M	0	4150000.00	Cs	G
210	M	0	4240000.00	Cs	G
211	M	11	4420000.00	Cs	G
212	M	0	4670000.00	Cs	G
213	M	0	4815000.00	Cs	A
214	M	0	5250000.00	Cs	A
215	M	0	6600000.00	Cs	A
216	M	0	6800000.00	Cs	A
217	M	0	7650000.00	Cs	G
218	M	0	7900000.00	Cs	G
219	M	0	7950000.00	Cs	G

220	M	0	8650000.00	Cs	A
221	M	0	9150000.00	Cs	G
222	M	0	9400000.00	Cs	G
223	M	17	10250000.00	Cs	G
224	M	0	10500000.00	Cs	G
225	M	0	11600000.00	Cs	G
226	M	0	12850000.00	Cs	G
227	M	0	14000000.00	Cs	G
228	M	0	14700000.00	Cs	G
229	M	0	14800000.00	Cs	G
230	M	0	16200000.00	Cs	A
231	M	0	17050000.00	Cs	G
232	M	0	18900000.00	Cs	G
233	M	0	18900000.00	Cs	G
234	M	0	19650000.00	Cs	G
235	M	30	20450000.00	Cs	G
236	M	114	22250000.00	Cs	G
237	M	0	24600000.00	Cs	G
238	M	12	24650000.00	Cs	G
239	M	0	28350000.00	Cs	G
240	M	35	28400000.00	Cs	G
241	M	0	29550000.00	Cs	G
242	M	0	34150000.00	Cs	A
243	M	0	34600000.00	Cs	G
244	M	21	39500000.00	Cs	G
245	M	0	55000000.00	Cs	A
246	M	0	57500000.00	Cs	A
247	M	0	60500000.00	Cs	A
248	M	0	63500000.00	Cs	G
249	M	17	73500000.00	Cs	A
250	M	0	294500000.00	Cs	G
251	M	0	320000000.00	Cs	G
252	M	0	341000000.00	Cs	G
253	M	21	377000000.00	Cs	G
254	M	0	454500000.00	Cs	G
255	M	43	482000000.00	Cs	G

References

Allen, M. I., Deslauriers, M., Andrews C. W., Tipples, G. A., Walters, K. A., Tyrell, D. L., Brown, N. and Condreay, L. D. (1998). Identification and characterization of mutations in hepatitis B virus resistant to lamivudine. Lamivudine Clinical Investigation Group. *Hepatology*. **27**: 1670-1677.

Baumert, T. F., Marrone, A., Vergalla, J. and Liang, T. J. (1998). Naturally Occurring Mutations Define a Novel Function of the Hepatitis B Virus Core Promoter in Core Protein Expression. J Virol. 72 (8): 6785-6795.

Beasley, R. P. and Hwang, L. Y. (1991). Overview on the epidemiology of hepatocellular carcinoma. In: *Viral Hepatitis and Liver Disease*. F.B. Hollinger, S.B. Lemon, H.S. Margolis (Eds.). Williams & Wilkins: Baltimore, 532-5.

Bosch, F. X., Ribes, J., Díaz, M. and Cléries, R. (2004). Primary liver cancer: worldwide incidence and trends. *Gastroenterology*. **127(5)**: S5-S16.

Boutwell, R. K. (1974). The function and mechanism of promoters of carcinogenesis. *CRC Crit Rev Toxicol.* **2**: 419-443.

Brechot, C., Gozuacik, D., Murakami, Y. and Paterlini-Brechot, P. (2000). Molecular bases for the development of hepatitis B virus (HBV)-related hepatocellular carcinoma (HCC). *Semin Cancer Biol.* **10**: 211-231.

Brechot, C., Pourcel, C., Louise, A., Rain, B. and Tiollais, P. (1980). Presence of integrated hepatitis B virus DNA sequences in cellular DNA of human hepatocellular carcinoma. *Nature.* **286**: 533-535.

Buckwold, V. E., Chen, M. and Ou, J. H. (1997). Interaction of transcription factors RFX1 and MIBP1 with the γmotif of the negative regulatory element of the hepatitis B virus core promoter. *Virology.* **227**: 515-518.

Buckwold, V. E., Xu, Z., Chen, M., Yen, T. S. B. and Ou, J. H. (1996). Effects of a naturally occurring mutation in the hepatitis B virus basal core promoter on precore gene expression and viral replication. *J. Virol.* **70**:5845–5851.

Cappel, R., DeCuyper, F. and Van Beers, D. (1977). E antigen and antibody, and chronic hepatitis. *J Infect Dis.* **136**:617-622.

Chan, H. L. and Sung, J. J. (2006). Hepatocellular carcinoma and hepatitis B virus. *Semin Liver Dis.* **26**: 153-161.

Chan, H. L., Hui, A. Y., Wong, M. L., Tse, A. M., Hung, L. C., Wong, V. W. and Sung, J. J. (2004). Genotype C hepatitis B virus infection is associated with an increased risk of hepatocellular carcinoma. *Gut.* **53(10)**:1494-8.

Chan, H. L., Tse, C. H., Mo, F., Koh, J., Wong, V. W., Wong, G. L., Chan, Lam, S., Yeo, W., Sung, J. J. and Mok, T. S. (2008). High viral load and hepatitis B virus subgenotype ce are associated with increased risk of hepatocellular carcinoma. *J Clin Oncol.* **26(2)**:177-82.

Chan, H. L., Tsui, S. K., Tse, C. H., Ng, E. Y., Au, T. C., Bartholomeusz, A., Leung,K. S., Locarnini, S. and Sung, J. J. (2005). Epidemiological and virological characteristics of 2 subgroups of hepatitis B virus genotype C. *J Infect Dis.* **191(12)**:2022-32.

Chen, M. and Ou, J. H. (1995). Cell type-dependent regulation of the activity of the negative regulatory element of the hepatitis B virus core promoter. *Virology.* **214**:198–206.

Cheng, Y., Seet, B. L., Ong, C. S., Wasser, S., Tan, T. M., Peter, F. J. and Lim, S. G. (2006). Are in vitro hepatitis B core promoter mutations important for clinical alterations in viral load? *Antiviral Res.* **69(3)**:142-51.

Chisari, F. V. and Ferrari, C. Viral Hepatitis. In: Nathanson N et al., eds. *Viral Pathogenesis*. Philadelphia, Lippincott – Raven, **1997**:745-778.

Chisari, F. V., Pinkert, C. A., Milich, D. R., Filippi, P., McLachlan, A., Palmiter, R. D. and Brinster, R. L. (1985). A transgenic mouse model of the chronic hepatitis B surface antigen carrier state. *Science.***230**: 1157-1160.

Cougot, D., Neuveut, C., Buendia, M. A. (2005). HBV induced carcinogenesis. *J Clin Virol.* **34** Suppl 1: S75-S78.

Das, K., Xiong, X., Yang, H., Westland, C. E., Gibbs, C. S., Sarafianos, S. G and Anold, E. (2001). Moleclar modeling and biochemical characterization reveal the mechanism of hepatitis B virus resistant to lamivudine (3TC) and emtricitabine (FTC). *J Virol.* **75**: 4771-4779.

Deguchi, M., Yamashita, N., Kagita, M., Asari, S., Iwatani, Y. and Tsuchida, T. (2004). Quantitation of hepatitis B surface antigen by an automated chemiluminescent microparticle immunoassay. *J Virol Methods* **115(2)**:217–222.

Dong, Q., Chan, H. L., Liu, Z., Chan, D. P., Zhang, B., Chen, Y., Kung, H., Sung, J. J. and He, M. (2008). A1762T/G1764A mutations of hepatitis B virus, associated with the increased risk of hepatocellular carcinoma, reduce basal core promoter activities. *BBRC.* **374**:773-776.

Elmore, L. W., Hancock, A. R., Chang, S. F., Wang, X. W., Chang, S., Callahan, C. P., Geller, D. A., Will, H. and Harris, C. C. (1997). Hepatitis B virus X protein and p53 tumor suppressor interactions in the modulation of apoptosis. *Proc Natl Acad Sci USA.* **94**:14707-14712.

Fang, Z. L., Sabin, C. A., Dong, B. Q., Wei, S. C., Chen, Q. Y., Fang, K. X., Yang, J. Y., Wang, X. Y. and Harrison, T. J. (2009). The association of HBV core promoter double mutations (A1762T and G1764A) with viral load differs between HBeAg positive and anti-HBe positive individuals: A longitudinal analysis. *J Hepatol.* **50**: 273-280.

Ferrari, C., Missale, G.,Boni, C. and Urbani, S. (2003). Immunopathogenesis of hepatitis B. *J Hepatol.* **39** Suppl 1: S36-S42.

Ganem, D. and Prince, A. M. (2004). Hepatitis B virus infection-natural history and clinical consequences. *N Engl J Med.* **350**: 1118-29.

Ganem, D. and Schneider, R. J. Hepadnaviridae: The Viruses and Their Replication. In: Knipe DM et al., eds. *Fields Virology*, 4th ed. Philadelphia, Lippincott Williams & Wilkins. **2001**:2923-2969.

Gerlach, K. K. and Schloemer, R. H. (1992). Hepatitis B virus C gene promoter is under negative regulation. *Virology*. **189(1)**:59-66.

Guidotti, L. G., Matzke, B., Pasquinelli, C., Shoenberger, J. M., Rogler, C. E. and Chisari, F. V. (1996). The hepatitis B virus (HBV) precore protein inhibits HBV replication in transgenic mice. *J Virol*. **70(10)**:7056-61.

Gunther, S., N. and Piwon, H. Will. (1998). Wild-type levels of pregenomic RNA and replication but reduced pre-C RNA and e-antigen synthesis of hepatitis B virus with C(1653) → T, A(1762) → T and G(1764) → A mutations in the core promoter. *J Gen Virol*. **79**:375-380.

Gunther, S., N. Piwon, A. Iwanska, R. Schilling, H. Meisel, H. and Will. (1996). Type, prevalence, and significance of core promoter/enhancer II mutations in hepatitis B viruses from immunosuppressed patients with severe liver disease. *J Virol*. **70**: 8318-8331.

Guo, W., Chen, M., Yen, T. S. and Ou, J. H. (1993). Hepatocyte-specific expression of the hepatitis B virus core promoter depends on both positive and negative regulation. *Mol Cell Bio.* **13**: 443-8.

Hildt, E., Munz, B., Saher, G., Reifenberg, K., Hofschneider, P. H. (2002). The PreS2 activator MHBs(t) of hepatitis B virus activates c-raf-1/Erk2 signaling in transgenic mice. *EMBO J.* **21**: 525-535.

Hollinger, F. B. and Liang, T. J. Hepatitis B Virus. In: Knipe DM et al. eds. *Fields Virology*, 4th ed. Philadelphia, Lippincott Williams & Wilkins, **2001:**2971-3036.

Horikawa, I. and Barrett, J. C. (2001). cis-Activation of the human telomerase gene (hTERT) by the hepatitis B virus genome. *J Natl Cancer Inst.* **93**: 1171-1173.

Hytiroglou, P. (2006). Theise ND. Telomerase activation in human Hepatocarcinogenesis. *Am J Gastroenterol* 2006; **101**: 839-841.

Iizuka, N., Tsunedomi, R., Tamesa, T., Okada, T., Sakamoto, K., Hamaguchi, T., Okabe. H. Y., Miyamoto, T., Uchimura, A., Hamamoto, Y. and Oka, M. (2006). Involvement of c-myc-regulated genes in hepatocellular carcinoma related to genotype-C hepatitis B virus. *J Cancer Res Clin Oncol.* **132**: 437-481.

Jammeh, S., Tavner, F., Watson, R., Thomas, H. C. and Karayiannis, P. (2008). Effect of basal core promoter and pre-core mutations on hepatitis B virus replication. *J Gen Virol.* **89**:901-909.

Jean-Michel, P. (2005). The concept of hepatitis B virus mutant escape. *J Clin Viral*. **34**: Suppli. 1. S125-129.

Kao, J. H., Chen, P. J., Lai, M. Y. and Chen, D. S. (2000). Hepatitis B genotypes correlates with clinical outcomes in patients with chronic hepatitis B. *Gastrolenterology*. **118**: 554-559.

Kao, J. H., Chen, P. J., Lai, M. Y. and Chen, D. S. (2003). Basal core promoter mutations of hepatitis B virus increase the risk for hepatocellular carcinoma in hepatitis B carriers. *Gastroenterology*. **124**: 327-334.

Koike, K., Tsutsumi, T., Fujie, H., Shintani, Y. and Kyoji, M. (2002). Molecular mechanism of viral hepatocarcinogenesis. *Oncology.* **62** Suppl 1: 29-37.

Kramvis, A. and Kew, M. C. (1999). The core promoter of hepatitis B virus. *J Viral Hepat* . **6**: 415-27.

Lai, C. L. and Locarnini, S. (2002). *Hepatitis B Virus.* Lodon: International Medical Press Ltd.

Lai, C. L. and Locarnini, S. (2008). *Hepatitis B Virus (2nd Edition).* Lodon: International Medical Press Ltd.

Lamberts, C., Nassal, M., Velhagen, I., Zentgraf, H. and Schroder, C. H. (1993). Precore-mediated inhibition of hepatitis B virus progeny DNA synthesis. J Virol. 67 (7): 3756-3762

Lanford, R. E. Notvall, L., Lee, H. and Beames, B.. (1997). Transcomplementation of nucleotide priming and reverse transcription between independently expressed TP and RT domains of the hepatitis B virus reverse transcriptase. *J Virol.* **71**: 2996-3004.

Laras, A., Koskinas, J.and Hadziyannis, S. J. (2002). In vivo suppression of precore mRNA synthesis is associated with mutations in the hepatitis B virus core promoter. *Virology.* **295**: 86-96.

Lee, S. G. and Rho, H. M. (2000). Transcriptional repression of the human p53 gene by hepatitis B viral X protein. *Oncogene.* **19**: 468-471.

Lei, J. H., Yang, X., Luo, H. Y., Wang, W. L. and Huang, L. (2006). Serum HBsAg concentration and HBV replication level in hepatitis B patients with positive serum HBsAg and HBeAg. *Zhong Nan Da Xue Xue Bao Yi Xue Ban.* **31(4)**:548–551.

Lesmana, L. A., Leung, N. W. Y. and Mahachai, V. (2006). Hepatitis B overview of the burden of disease in the Asia-Pacific region. *Liver International.* **26**: 3-10.

Liu, T. T., Fang, Y., Xiong, H., Chen, T. Y., Ni, Z. P., Luo, J. F., Zhao, N. Q. and Shen, X. Z. (2008). A case-control study of the relationship between hepatitis B virus DNA level and risk of hepatocellular carcinoma in Qidong, China. *World J Gastroenterol.* **14(19)**:3059-3063.

Llovet, J. M., Burroughs, A. and Bruix, J. (2003). Hepatocellular carcinoma. *Lancet.* **362(9399)**: 1907-1917.

Lo, W. Y. and Ting L. P. (1994). Repression of enhancer II activity by a negative regulatory element in the hepatitis B virus genome. *J Virol.* **68**: 1758–1764.

Mahoney, F. J. and Kane, M. Hepatitis B Vaccine. In: Plotkin SA, Orenstein WA, eds. *Vaccines*, 3rd ed. Philadelphia, W.B. Saunders Company, **1999**:158-182.

Maynard, J. E. (1990). Hepatitis B: Global importance and need for control Vaccine, **8(Suppl.)**: S18

Maynard, J. E. Control of hepatitis B by immunization: Global perspectives. In: *Viral Hepatitis and Liver Disease*. Grunne & Stratton: New York, **1988**: 967-9

McGlynn, K. A. and London, W. T. (2005). Epidemiology and natural history of hepatocellular carcinoma. *Best Pract Res Clin Gastroenterol.* **19(1)**: 3-23.

Mendenhall, C. L., Seeff, L., Diehl, A. M., Ghosn, S. J., French, S. W. and Gartside, P. S. (1991). Antibodies to hepatitis B virus and hepatitis C virus in alcoholic hepatitis and cirrhosis: their prevalence and clinical relevance. The VA Cooperative Study Group (No. 119). *Hepatology.* **14(4)**: 581–589.

Mendy, M. E., Kaye, S., van der Sande, M., Rayco-Solon, P., Waight, P. A.,

Shipton, D. D., Awi, D., Snell, P., Whittle, H. and McConkey, S. J. (2006). Application of real-time PCR to quantify hepatitis B virus DNA in chronic carriers in The Gambia. *Virol J.* **3:**23.

Miyakawa, Y. and Mzokami, M. (2003). Classifying hepatitis B virus genotypes. *Intervirology* **46**: 329-328.

Moolla, N., Kew, M. and Arbuthnot, P. (2002). Regulatory elements of hepatitis B virus transcription. *J Viral Hepat.* **9**: 323-31.

Noboru, S., Yasuhito, T. and Kioaki, I. (2007). Influence of hepatitis B virus X, core promoter mutations on hepatocellular carcinoma among patients with subgenotype C2. *JCM* **10**: 3191-3197.

Nordenfelt, E. and L. Kjellen. 1975. Dane particles, DNA polymerase, and e-antigen in two different categories of hepatitis B antigen carriers. Intervirology. 5:225-232.

Orito, E., Ichida, T., Sakugawa, H., Sata, M., Horiike, N., Hino, K., Okita, K., Okanoue, T., Iino, S., Tanaka, E., Suzuki, K., Watanabe, H., Hige, S. and Mizokami, M. (2001). Geographic distribution of hepatitis B virus (HBV) genotype in patients with chronic HBV infection in Japan. *Hepatology.* **34**: 590-4.

Osama, Y. and Makoto, A. (2006). Molecular biology of hepatitis B virus: effect of nucleotide substitutions on the clinical features of chronic hepatitis B. *Mef Mol Morphol* . **39**: 113-120.

Parekh, S., Zoulium, F., Ahn, S. H. Tsai,, A., Li, J. and Kawai, S. (2003). Genome replication, viron secretion, and e antigen expression of naturally occurring hepatitis B virus core promoter mutants. *J Virol.* **77**:6601-6612.

Pares, A., Barrera, J. M., Caballeria, J., Ercilla, G., Bruguera, M. and Caballeria, L. (1990). Hepatitis C virus antibodies in chronic alcoholic patients: association with severity of liver injury. *Hepatology.* **12(6)**:1295–1299.

Park, G. T., Yi, Y. W., Choi, C. Y. and Rho, H. M. (1997). A negative regulatory element and its binding protein in the upstream of enhancer II of hepatitis B virus. *DNA Cell Biol.* **16**: 1459–1465.

Paterlini-Brechot, P., Saigo, K., Murakami, Y., Chami, M., Gozuacik, D., Mugnier, C., Lagorce, D. and Brechot, C. (2003). Hepatitis B virus-related insertional mutagenesis occurs frequently in human liver cancers and recurrently targets human telomerase gene. *Oncogene.* **22**: 3911-3916.

Paterlini-Brechot, P., Saigo, K., Murakami, Y., Chami, M., Gozuacik, D., Mugnier, C., Lagorce, D. and Brechot, C. (2003). Hepatitis B virus-related insertional mutagenesis occurs frequently in human liver cancers and recurrently targets human telomerase gene. *Oncogene.* **22**: 3911-3916.

Pult, I., Chouard, T., Wieland, S., Klemenz, R., Yaniv, M. and Blum, H. E. (1997). A hepatitis B virus mutant with a new hepatocyte nuclear factor binding site emerging in transplant-transmitted fulminant hepatitis B. *Hepatology.* **25**: 1507-1515.

Rehermann, B. and Nascimbeni, M. (2005). Immunology of hepatitis B virus and hepatitis C infection. *Nature Reviews Immunology* **5:**215-229.

Reinhold, W., Emens, L., Itkes, A., Blake, M., Ichiro, I. and Zajac-kaye, M. (1995). The myc intron-binding polypeptide associates with RFX1 in vivo and binds to the major histocompatibility complex class II promoter region, to the hepatitis B virus enhancer, and to regulatory regions of several distinct viral genes. *Mol Cell Bio.* **15(6)**:3041-3048.

Robinson, W. S. Hepatitis B Virus and hepatitis D virus. In: Mandell GL, Bennett JE, Dolin R, eds. *Principles and Practice of Infectious Diseases*, 4th ed. New York, Churchill Livingstone, **1995**:1406-1439.

Sawada, M., Takada, Takase, A. and Takada, S., N. (1993). Effects of alcohol on the replication of hepatitis C virus. *Alcohol Alcohol Suppl* . **1B**:85–90.

Scaglioni, P. P., Melegari, M. and Wands, J. R. (1997). Posttranscriptional regulation of hepatitis B virus replication by the precore protein. J Virol. 71(1): 345-353.

Sheldon, J., Rodes, B., Zoulim, F., Bartholomeusz and Soriano, V. (2005). Mutations affecting the replication capacity of the hepatitis B virus. *J Virol Hepatitis*. **13**: 427-434.

Shinkai, N., Tanaka, Y., Ito, K., Mukaide, M., Hasegawa, I., Asahina, Y., Izumi, N., Yatsuhashi, Y., Orito, E., Joh, T. and Mizokami, M. (2007). Influence of hepatitis B virus X and core promoter mutations on hepatocellular carcinoma among patients infected with subgenotype C2. *J Clin Microbiol.* **45**:3191-3197.

Staib, F., Hussain, S. P., Hofseth, L. J., Wang, X. W. and Harris, C. C. (2003). TP53 and liver carcinogenesis. *Hum Mutat.* **21**: 201-216.

Sumi, H., Yokosuka, O., Seki, N., Arai, M., Imazeki, F., Kurihara, T., Kanda, T., Fukai, K., Kato, M. and Sasho, H. (2003). Influence of hepatitis B virus genotypes on the progression of chronic type B liver disease. *Hepatology.* **37**: 19-26.

Sun, C. T., Lo, W. Y., Wang, I. H., Lo, Y. H., Shiou, S. R., Lai, C. K. and Ting, L. P. (2001). Transcription repression of human hepatitis B virus genes by negative regulatory element-binding protein/SON, *J. Biol. Chem.* **276**: 24059–24067.

Sung, J. J. Y., Tsui, S. K. W., Tse, C. H., Ng, Y. E. T., Leung, K. S., Lee, K. H., Mok, T. S. K., Bartholomeusz, A., Au, T. C. C., Zhang, M. D. and Chan, H. L. Y.. 2008. Genotype-specific genomic markers associated with primary hepatomas, based on complete genomic sequencing of hepatitis B virus. *J Virol.* **82(7)**:3604-11.

Tacke, F., Gehrke, C., Luedde, T., Heim, A., Manns, M. P. and Trautwein, C. (2004). Basal core promoter and precore mutations in the hepatitis B virus genome enhance replication efficacy of Lamivudine-resistant mutants. *J Virol.* 78(16):8524-35.

Takahashi, K., Akahane, Y., Hino, K., Ohta, Y. and Mishiro, S. (1998). Hepatitis B virus genomic sequence in the circulation of hepatocellular carcinoma patients: comparative analysis of 40 full-length isolates. *Virology.* **143(12)**: 2313-26.

Takahashi, K., Kishimoto, S., Ohori, K., Yoshizawa, H., Machida, A., Ohnuma, H., Tsuda, F., Munekata, E., Miyakawa, Y., and Mayumi, M. (1991). Molecular heterogeneity of e antigen polypeptides in serafrom carriers of hepatitis B virus. *J. Immunol.* 147, 3156–3160.

Thorgeirsson, S. S. and Grisham, J. W. (2002). Molecular pathogenesis of human hepatocellular carcinoma. *Nat Genet.* **31**: 339-346.

Ueda, H., Ullrich, S. J., Gangemi, J. D., Kappel, C. A., Ngo, L., Feitelson, M. A. and Jay, G. (1995). Functional inactivation but not structural mutation of p53 causes liver cancer. *Nat Genet.* **9**: 41-47.

Visvanathan, K. and Lewin, S. R. (2006). Immunopathogenesis: role of innate and adaptive immune responses. *Semin Liver Dis.* **26**: 104-115.

Wang, J., Chenivesse, X., Henglein, B. and Brechot, C. (1990). Hepatitis B virus integration in a cyclin A gene in a hepatocellular carcinoma. *Nature.* **343**: 555-557.

Waris, G. and Siddiqui, A. (2003). Regulatory mechanisms of viral hepatitis B and C. *J Biosci.***28**: 311-321.

WHO report 2006, Genava, Switzerland.

Will, H., Reiser, W., Weimer, T., Pfaff, E., Buscher, M., Sprengle, R.,Cattaneo, R., and Schaller, H. (1987). Replication strategy of humanhepatitis B virus. *J. Virol.* **61**, 904–911.

Wong, V. W. S., Chan, S. L., Mo, F., Chan, T. C., Loong, H. H. F., Wong, G. L. H., Lai, Y. Y. N., Chan, A. T. C., Sung, J. J. Y., Yeo, W., Chan, H. L. Y. and Mok, T. S. K. (2010). Clinical Scoring System to Predict Hepatocellular Carcinoma in Chronic Hepatitis B Carriers. *J Clin Oncol.* **28(10)**:1660-5.

Xie, Y., Zhao, H., Dai, W. S. and Xu, D. Z. (2003). HBV DNA level and antigen concentration in evaluating liver damage of patients with chronic hepatitis B. *Hepatobiliary Pancreat Dis Int.* **2(3)**:418–422.

Yaginuma, K., Shirakata, Y., Kobayashi, M., and Koike, K. (1987). Hepatitis B virus (HBV) particles are produced in a cell culture system by transient expression of transfected HBVDNA. *Proc. Natl. Acad. Sci. USA* 84, 2678–2682.

Yeung, Y. P., Lo, C. M., Liu, C. L., Wong, B. C., Fan, S. T. and Wong, J. (2005). Natural history of untreated nonsurgical hepatocellular carcinoma. *Am J Gastroenterol.* **100(9)**: 1995-2004.

Yu, X., and J. E. Mertz. 1996. Promoters for synthesis of the pre-C and pregenomic mRNAs of human hepatitis B virus are genetically distinct and differentially regulated. J. Virol. **70:**8719–8726.

Yuh, C. H., Chang, Y. L. and Ting, L. P. (1992). Transcription regulation of precore and pregenomic RNAs of hepatitis B virus. J Virol. **66(7)**:4073-84.

Zhang, X. Y., Jabrane-Ferrat, N., Asiedu, C. K., Samac, S., Peterlin, B. M. and Ehrlich, M. (1993). The major histocompatibility complex class II promoter-binding protein RFX (NF-X) is a methylated DNA-binding protein. *Mol Cell Bio*. **13(11)**:6810-6818.

Zhang, X., Zhang, H. and Ye, L. (2006). Effects of hepatitis B virus X protein on the development of liver cancer. *J Lab Clin Med.* **147**: 58-66.

Zheng, Y., Li, J. and Ou, J. H. (2004). Regulation of hepatitis B virus core promoter by transcription factors HNF1 and HNF4 and the viral X protein. *J Virol*. **78**: 6908-6914.