Sequence Variations of Full-Length Hepatitis B Virus Genomes in Chinese Patients with HBsAg-Negative Hepatitis B Infection

Fung-Yu Huang[1]◯, Danny Ka-Ho Wong[1,2]◯, Wai-Kay Seto[1], An-Ye Zhang[1], Cheuk-Kwong Lee[3], Che-Kit Lin[3], James Fung[1,2], Ching-Lung Lai[1,2], Man-Fung Yuen[1,2]*

1 Department of Medicine, The University of Hong Kong, Queen Mary Hospital, Hong Kong, China, 2 State Key Laboratory for Liver Research, The University of Hong Kong, Queen Mary Hospital, Hong Kong, China, 3 Hong Kong Red Cross Blood Transfusion Service, Hospital Authority, Hong Kong, China

Abstract

Background: The underlying mechanism of HBsAg-negative hepatitis B virus (HBV) infection is notoriously difficult to elucidate because of the extremely low DNA levels which define the condition. We used a highly efficient amplification method to overcome this obstacle and achieved our aim which was to identify specific mutations or sequence variations associated with this entity.

Methods: A total of 185 sera and 60 liver biopsies from HBsAg-negative, HBV DNA-positive subjects or known chronic hepatitis B (CHB) subjects with HBsAg seroclearance were amplified by rolling circle amplification followed by full-length HBV genome sequencing. Eleven HBsAg-positive CHB subjects were included as controls. The effects of pivotal mutations identified on regulatory regions on promoter activities were analyzed.

Results: 22 and 11 full-length HBV genomes were amplified from HBsAg-negative and control subjects respectively. HBV genotype C was the dominant strain. A higher mutation frequency was observed in HBsAg-negative subjects than controls, irrespective of genotype. The nucleotide diversity over the entire HBV genome was significantly higher in HBsAg-negative subjects compared with controls ($p = 0.008$) and compared with 49 reference sequences from CHB patients ($p = 0.025$). In addition, HBsAg-negative subjects had significantly higher amino acid substitutions in the four viral genes than controls (all $p < 0.001$). Many mutations were uniquely found in HBsAg-negative subjects, including deletions in promoter regions (13.6%), abolishment of pre-S2/S start codon (18.2%), disruption of pre-S2/S mRNA splicing site (4.5%), nucleotide duplications (9.1%), and missense mutations in "α" determinant region, contributing to defects in HBsAg production.

Conclusions: These data suggest an accumulation of multiple mutations constraining viral transcriptional activities contribute to HBsAg-negativity in HBV infection.

Introduction

Hepatitis B virus (HBV) infection leads to a wide spectrum of liver injuries, ranging from acute self-limiting infection and fulminant hepatitis to chronic hepatitis, liver cirrhosis and hepatocellular carcinoma (HCC) [1]. The viral genome contains four partially overlapping open reading frames (ORFs) (pre-S/S, pre-C/C, P and X) [2]. HBV infection is usually diagnosed when the circulating HBV surface antigen (HBsAg) is detected. Spontaneous loss of HBsAg is a rare event in chronic hepatitis B (CHB) infection [3–6]. Loss of serum HBsAg is also observed in patients receiving treatment for CHB with interferon and nucleoside/nucleotide analogues [7,8].

Although the continued presence of detectable HBsAg in serum is a prominent feature of HBsAg-positive CHB infection or overt CHB infection, HBsAg seroclearance may not signify eradication of HBV. The introduction of sensitive HBV DNA detection tests has revealed the existence of HBsAg-negative HBV infection. This new entity of HBV infection is defined as "the presence of viral DNA in the liver (with detectable or undetectable HBV DNA in the serum) of individuals testing negative for the HBsAg" [9,10]. Some of these subjects may be people with or without prior medical history of HBV infection can develop to the occult phase [11]. All these scenarios are collectively termed occult hepatitis B infection (OBI) [10].

Clinical evidence reveals that there is generally no reduction in the risk for HCC in CHB patients with HBsAg seroclearance

compared with those who are persistently positive for HBsAg [12–14]. The virological features and the mechanisms leading to OBI are still unclear, although viral mutations may be one of the significant factors. Many attempts have been done to analyze HBV genome sequences amplified from OBI patients. Most studies focused on searching for mutations on a fragment of HBV genome (mainly the region coding for HBsAg). This is due to the extremely low viral DNA levels in blood or liver tissue in these patients rendering amplification difficult. Mutations have been found in various regions of the HBV genome and may be associated with OBI: (1) mutations in the surface protein affecting antigen detection [15–17]; (2) deletions in the pre-S1 region that impair HBV packaging [18,19]; (3) various mutations in viral regulatory regions that cause decreased in HBsAg expression and/or viral replication [20,21]; and (4) mutations affecting post-translational protein production [20,22]. In addition to these potential mechanisms, epigenetic modifications of the HBV genome are also suggested to play a significant role in OBI [23]. However, these studies mainly evaluated viral coding sequences, in particular, the different surface proteins. Owing to the complex organization of viral genome and also the efficient regulation, it is important to study mutations in both gene coding regions as well as the regulatory regions. It remains elusive as to which mutations and variations can lead to OBI. Thus, comprehensive whole genome analyses are still needed.

In the present study, we adopted a highly efficient amplification method, rolling circle amplification (RCA) to amplify full-length HBV genome from sera and liver biopsies of OBI subjects. We primarily aimed at analyzing sequence variations of complete HBV genome in subjects with OBI. The secondary aim was to analyze the effect of pivotal mutations identified on regulatory region(s) on promoter activities of the viral genes.

Materials and Methods

Patients and Samples

A total of 185 serum samples (156 from CHB patients with HBsAg seroclearance; 29 from HBsAg-negative patients with detectable HBV DNA at the first presentation) and 60 liver biopsies (21 from CHB patients with HBsAg seroclearance; 39 from HBsAg-negative patients with detectable HBV DNA at the first presentation) were recruited. Another 11 serum samples were collected from HBsAg-positive treatment-naïve CHB patients as controls. These subjects were first identified by testing their blood samples for HBsAg (Abbott Prism, Abbott Laboratories, Abbott Park IL) and by the nucleic acid amplification test (NAT) for HBV DNA (Procleix Trigis system, Novartis Diagnostics, Emeryville, CA; 95% detection limit, 12.2 IU/ml). Serum HBV DNA levels were measured using the COBAS TaqMan HBV Monitor Test (Roche Diagnostics, Branchburg, NJ), with a lower limit of detection of 20 IU/ml (100 copies/ml). HBV DNA levels in the liver biopsies were measured by real-time PCR using the Artus HBV RG assay (Qiagen), which has a linear range of detection of 1.1 IU/ml to $>4 \times 10^9$ IU/ml. The present study was approved by the Institutional Review Board, the University of Hong Kong, Hong Kong. Written informed consent was obtained from all patients.

Rolling Circle Amplification and Sequencing of Full-length HBV Genome

RCA, a powerful PCR-based technique for the amplification of low viral load circular DNA, has been previously described [24]. Briefly, HBV plus-strand DNA inside the virions in the serum was completed using the endogenous HBV polymerase before extraction using the QIAamp MinElute Virus Vacuum Kit (Qiagen). Nicks in the HBV DNA strands were then completed by ligase (Epicentre, Madison, WI). For the liver biopsy samples, total DNA was extracted using the QIAamp Allprep Kit (Qiagen). The HBV DNA samples were subjected to RCA using eight primers spanning the whole HBV genome (Table S1) [24]. Single genome-length HBV DNA was either retrieved directly by *Spe*I digestion (New England Biolabs) or further amplified using the Expand High Fidelity PCR system (Roche, Mannheim, Germany) [25]. This amplified HBV DNA was then purified and sequenced using sequencing primers that cover the whole viral genome (Table S1). Sensitivity and conditions for RCA reaction were tested against samples with known copies of viral DNA (quantified by the Artus real-time PCR kit; Qiagen).

Nucleotide Sequence and Amino Acid Analysis

Full-length HBV sequences were assembled using the CLC Main Workbench 6.6.2 (CLC Bio, Katrinebjerg, Denmark). HBV genotype was determined using the NCBI genotyping tool and phylogenetic analysis. HBV sequences from OBI subjects were compared with either sequences from the control subjects or genotype-matched HBV sequences of Chinese CHB patients retrieved from the GenBank [26]. Analysis of nucleotide and amino acid diversity (d), as well as phylogenetic analysis, was performed using the MEGA version 5 software [27]. The program SimPlot [28] was used to search for evidence of recombination. Amino acid mutation frequency was defined as the number of amino acid variations per total number of amino acid residues within a particular genomic region. Prediction of RNA secondary structure was performed using the MFOLD software [29].

In-vitro Analysis on the Effects of Genetic Mutations on Specific Regulatory Regions on HBV Promoter Activities

HBV promoter regions from selected OBI cases and wilde-type controls were amplified using primers containing *Kpn*I and *Nhe*I restriction sites. The *Kpn*I- and *Nhe*I-digested PCR products, containing different mutated HBV promoter regions, were cloned into the reporter vector pGL3-Basic. Positive constructs were confirmed and co-transfected with pRL-TK *Renilla* reporter plasmid into Huh-7 cells using Lipofectamine 2000 (Invitrogen). After 24 hours of incubation, the cells were assayed for luciferase activity using the Dual-Luciferase Reporter Assay System (Promega, Madison, WI). Promoter activities were expressed as a ratio of firefly luciferase to *Renilla* luciferase luminescence. Results were expressed in arbitrary units (AU).

Statistical Analyses

Differences between categorical variables were analyzed using the Fisher's exact test or Chi-square test. For continuous variables, the Student's t-test was used. All statistical analysis was done using GraphPad Prism 5.0 (GraphPad Software, Inc. San Diego, CA). Data are expressed as percentage or mean ± SD. A p-value of < 0.05 was considered to be statistically significant.

Results

Rolling Circle Amplification for HBV Genomes from OBI

By using known copies of HBV DNA isolated from CHB patients, we demonstrated RCA could amplify down to 15 copies/reaction (Figure 1). Full-length HBV genomes were successfully amplified from 18/60 (30%) liver biopsies and 4/185 (2.2%) serum samples from OBI subjects using this RCA method. Complete HBV genomes were also amplified from all 11 controls. Seventeen out of these 18 OBI subjects had detectable intrahe-
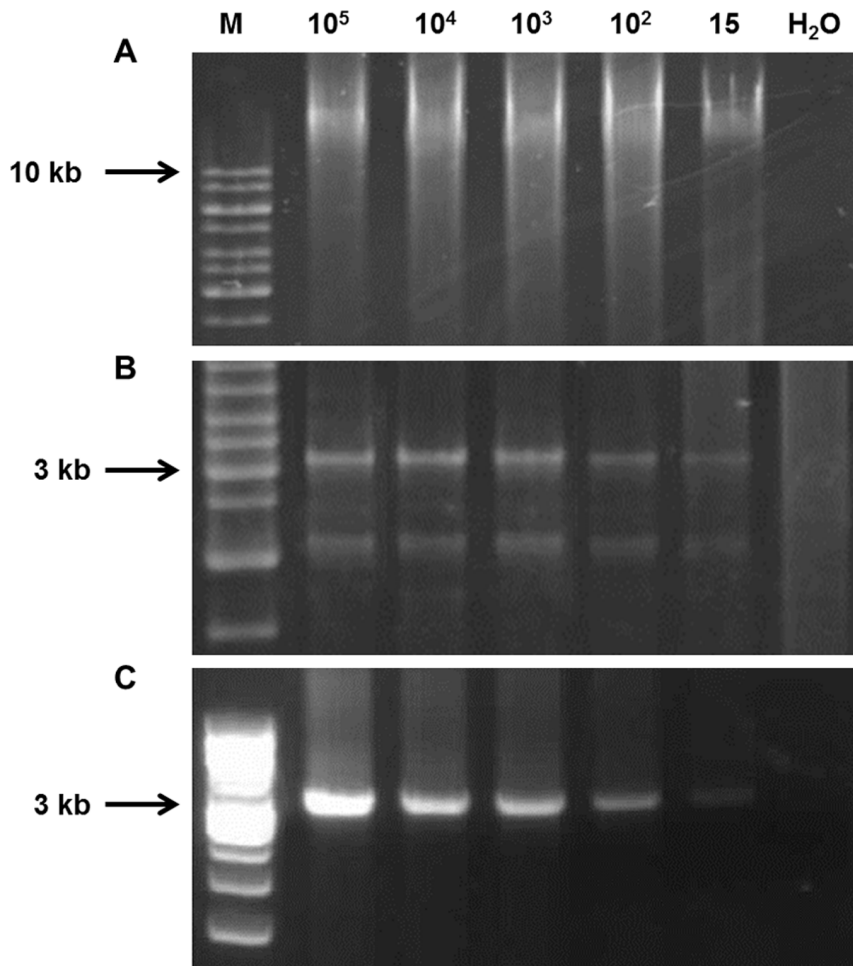
Figure 1. Rolling circle amplification (RCA) of full-length HBV genome. The number of initial HBV DNA template present in each RCA reaction is shown at the top. M, molecular weight marker; H2O, negative control. (A) The high-molecular-weight raw RCA products containing multiple copies of the initial HBV template. (B) The full-length HBV genomes recovered by restriction enzyme, SpeI digestion. (C) Full-length HBV genome amplified by using the RCA products as PCR template.
doi:10.1371/journal.pone.0099028.g001

patic HBV DNA levels (median: 4.07 copies/cell; range: 0.07–15.34 copies/cell). HBV viral load in the 4 serum samples were all under the detection limit of COBAS assay of 20 IU/ml. Full-length HBV genomes were also amplified from 11 HBsAg-positive controls (median: 3.51×10^5 IU/ml; range: $100–7.34 \times 10^7$ IU/ml).

Demographic Data and Phylogenetic Analysis

Demographic data of the study subjects are summarized in Table 1. Genotype C was dominant in both OBI group (17/22; 77.3%; 11 with subtype C1, 4 with subtype C2 and 2 outliers) and controls (9/11; 81.8%; all subtype C1). Genotype B was detected in the remaining OBI subjects (5/22; 22.7%; all subtype B2) and controls (2/11; 18.2%; 1 with subtype B2 and 1 outlier) (Figure S1). Evidence of intergenotypic recombination between genotypes B/C was detected in these 3 outliers but only constituted less than 20% of the major parental genotype. The full-length HBV genomes and multiple sub-genomic regions reproduce the same genotyping results, these 3 outliers (2 occult samples and 1 control) were clustered into genotype C and B respectively. Of note, phylogenetic analysis based on the full-length HBV sequences or

the 4 individual ORFs cannot distinctly separate OBI from CHB controls (data not shown).

Comparison of HBV Genomic Diversity between OBI and Control Subjects

The nucleotide and amino acid diversity were evaluated with respective to their genotypes. The nucleotide diversity over the entire HBV genome with genotype C was significantly higher in the 17 subjects with OBI when compared with the 9 control subjects (diversity, $d = 0.040 \pm 0.002$ vs. 0.026 ± 0.002, $p = 0.008$). The nucleotide diversity in these 17 OBI subjects was also significantly higher than that in 49 genotype C reference sequences retrieved from Chinese CHB patients [26] ($d = 0.04 \pm 0.002$ vs. $d = 0.030 \pm 0.001$, $p = 0.025$). There was no significant difference in the nucleotide diversity between OBI and control subjects with genotype B ($d = 0.021 \pm 0.002$ vs. $d = 0.016 \pm 0.002$, $p = 0.279$), which may be related to limited number of tested subjects (n = 5 and n = 2, respectively). Similarly, the nucleotide diversity between OBI subjects and 5 genotype B reference sequences retrieved from Chinese CHB patients showed no significant difference ($d = 0.016 \pm 0.002$ vs. $d = 0.015 \pm 0.002$, $p = 0.792$) [30] (Table 2). In terms of amino acid diversity, there

Table 1. Clinical characteristics of the study subjects with full-length HBV genome amplified by rolling circle amplification (RCA).

	OBI group	Control group	P-value
	(n = 22)	(n = 11)	
Age			
Mean ± SEM	49.18±6.16	54.91±16.54	0.156
Gender			
Male : Female	13: 9	5:6	0.488
HBV genotype (%)			
B	5 (22.7%)	2 (18.2%)	
C	17 (77.3%)	9 (81.8%)	0.763
Mutation frequency in the HBV key regulatory elements			
Enhancer I	17 (77.3%)	6 (54.5%)	0.240
Enhancer II	7 (31.8%)	3 (27.3%)	1.000
DR 1	1 (4.5%)	0 (0%)	n.a.
DR 2	0 (0%)	0 (0%)	n.a.
Pre-S1 promoter	11 (50.0%)	7 (63.6%)	0.712
Pre-S2 promoter	14 (63.6%)	6 (54.5%)	0.714
Core Promoter	22 (100%)	10 (91.0%)	0.333
X Promoter	12 (54.5%)	5 (45.5%)	0.721

(OBI, occult hepatitis B infection; n.a.: not applicable; SEM: standard error of the mean; DR: direct repeat).
doi:10.1371/journal.pone.0099028.t001

was no significant difference between OBI and control sequences for both genotypes (data not shown).

Further analysis on the nucleotide and amino acid diversity on the 4 individual HBV ORFs indicated OBI cases presented slightly higher nucleotide and amino acid diversity than control cases (Table 2). For genotype C cases, the nucleotide diversity in the pre-S1, pre-C and P ORFs was significantly higher in OBI than control cases (p = 0.048, p = 0.047 and p = 0.032, respectivel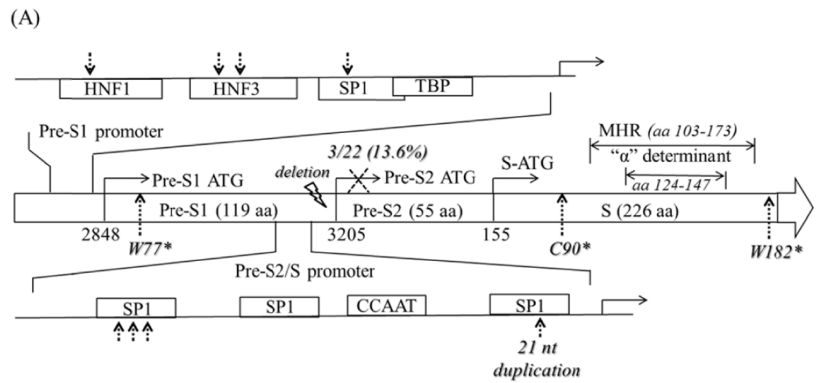y). For genotype B cases, OBI subjects had a significantly higher nucleotide diversity in the pre-C/C and C regions (p = 0.045 and p = 0.031, respectively) (Table 2), and in the RNase H region (p = 0.049) of the P ORF than control subjects (data not shown). Furthermore, the difference in the nucleotide diversity between OBI and control groups with genotype C in both pre-S/S and pre-C/C regions also showed a similar trend (p = 0.069 and p = 0.059, respectively) (Table 2). There was no significant difference on the amino acid diversity between OBI and control sequences on the 4 ORFs in both genotypes (data not shown). In summary, OBI

Table 2. Comparison of the nucleotide (nt.) diversity (± SEM) on the HBV genome of occult hepatitis B and control subjects.

HBV regions	Genotype B			Genotype C		
	OBI	Control	P-value	OBI	Control	P-value
Pre-S/S	0.008±0.002	0.006±0.002	0.490	0.030±0.003	0.018±0.003	0.069
Pre-S1	0.009±0.003	0.007±0.003	0.833	0.053±0.007	0.022±0.005	*0.048**
Pre-S2	0.011±0.005	0.002±0.002	0.320	0.032±0.008	0.024±0.009	0.603
S	0.007±0.002	0.006±0.002	0.706	0.020±0.003	0.015±0.003	0.461
Pre-C/Core	0.028±0.005	0.009±0.003	*0.045**	0.041±0.004	0.024±0.004	0.059
Pre-C	0.017±0.010	0.018±0.011	0.924	0.043±0.013	0.014±0.008	*0.047**
Core	0.030±0.005	0.008±0.003	*0.031**	0.041±0.004	0.025±0.004	0.099
P	0.013±0.002	0.014±0.002	0.911	0.039±0.003	0.026±0.002	*0.032**
X	0.016±0.004	0.036±0.006	0.090	0.027±0.004	0.018±0.004	0.267
Pre-S1 promoter	0.027±0.005	0.011±0.003	0.110	0.055±0.006	0.034±0.005	0.097
Pre-S2/S promoter	0.010±0.004	0.008±0.003	0.684	0.052±0.007	0.020±0.005	*0.037**
Full-length HBV	0.021±0.003	0.016±0.002	0.279	0.040±0.002	0.026±0.002	**0.008**
Δ Full-length HBV	0.016±0.002	0.015±0.002	0.792	0.040±0.002	0.030±0.002	*0.025**

ΔFull-length HBV: OBI sequence compared with 5 genotype B and 49 genotype C reference sequence obtained from Chinese patients with chronic HBV infection. *A statistical difference with p<0.05.
doi:10.1371/journal.pone.0099028.t002

Figure 2. Schematic representation showing types and positions of mutations affecting the HBV pre-S/S region. (A) Binding sites in the pre-S1 and pre-S2/S promoters for the liver-specific transcription factors (HNF1 and HNF3), Sp1, TATA binding protein (TBP) and the CCAAT box. Distinct mutations in TF binding sites (broken arrows), premature termination of surface protein synthesis (*), nucleotide duplication (broken arrows), deletion (a flash) and abolishment of pre-S2/S start codon (X) were indicated. (B) Amino acid changes in the pre-S1, pre-S2 and S genes. (Position with more than one amino acid changes was bracketed. Amino acid changes in "α" determinant region were underlined).
doi:10.1371/journal.pone.0099028.g002

cases had a higher nucleotide diversity than control cases, and these changes were scattered over the entire viral genome and may not lead to coding changes. This was evidenced by the comparable amino acid diversity between the two groups.
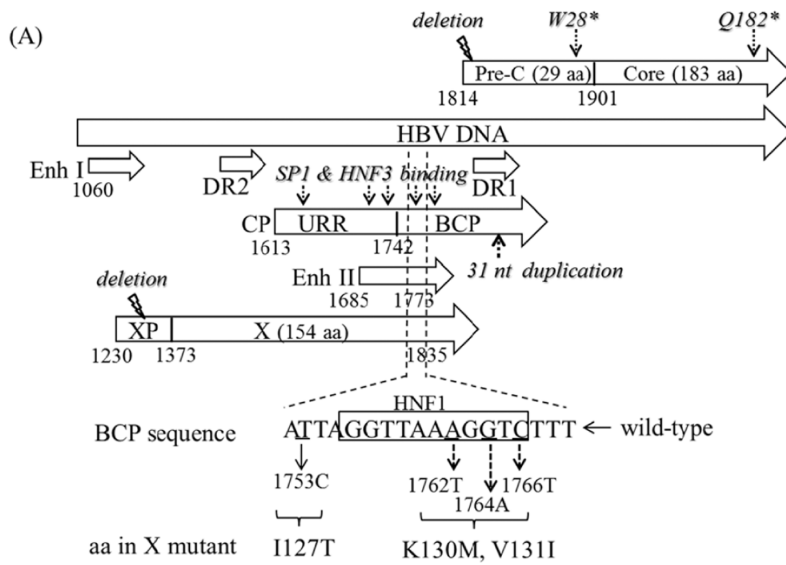
Mutational Analysis on the pre-S/S Coding Region

Details of the mutations observed in subjects with OBI and CHB infection were illustrated in Table S2. The total number of amino acid variations over the entire PreS/S region was significantly higher in OBI than control cases [89/400 (22.2%) vs. 33/400 (8.25%), p<0.0001] (Figure 2). The number of amino acid substitutions in individual pre-S1, pre-S2 and S regions were also significantly higher in OBI than control cases [pre-S1:21/119 (17.6%) vs. 3/119 (2.5%), p<0.0001; pre-S2:20/55 (36.4%) vs. 7/55 (12.7%), p<0.001; S: 48/226 (21.2%) vs. 23/226 (10.2%), p<0.001], respectively (Figure 2B). Some mutation patterns were uniquely observed in OBI cases in pre-S/S ORF, including sequence deletions (3/22, 13.6%) and abolishment of the pre-S2 start codon (ATG) by a point mutation (3/22, 13.6%). The analysis of S ORF showed that the clinically important amino acid substitutions were mainly located in the major hydrophilic region (MHR) (residues 103–173). These included I126S, I126N, Q129N, T131N, M133T, G145A and A159V. Except for A159V, all the other 6 amino acid substitutions in MHR reside in the "α" determinant region (residues 124–147) of the HBsAg, and were found more frequently in OBI than control cases [5/24 (20.8%) vs. 2/24 (8.3%), p = 0.416]. Among them, I126S, Q129N

and M133T are associated with diagnostic problems [15,31,32], whereas G145A is known as vaccine escape mutant and associated with OBI [33].

Mutational Analysis on the pre-C/C Coding Regions

The number of amino acid substitutions in OBI cases was significantly higher than control cases in both pre-C and core ORFs (pre-C: 18/29 (62.1%) vs. 3/29 (10.3%); core: 99/183 (54.1%) vs. 37/183 (20.2%), both with p<0.0001) (Figure 3A, 3B and Table S2). Within the pre-C region, the G1896A mutation introducing a stop signal at codon 28 (W28*), which abolishes HBeAg production was found in 7/22 occult cases and 1/11 control cases (31.8% vs. 9.1%, p = 0.2176) (Figure 3B). Mutations in core region were mainly presented within different antigen immunogenic epitopes. As shown in Figure 3B, OBI cases had a higher number of amino acid substitutions than control cases within the CD8+ CTL-epitope (residues 18–27) [6/10 (60%) vs. 1/10 (10%), p = 0.0573] and CD4+ T-cell epitope (residues 48–69) [9/22 (40.9%) vs. 3/22 (13.6%), p = 0.0883]. Similarly, multiple amino acid substitutions were also detected within the HBcAg domain (residues 80–120). In particular, 9 amino acid changes (S74G, E77D, E77Q, P79Q, A80T, E83D, L84A, S87G and S87R) within the HBc/e1 epitope (residues 75–90) were solely detected in OBI cases (56.3% vs. 0%, p<0.015), and that was known to reduce both HBe and HBc antigenicity [34]. In addition, amino acid substitutions within the HBc/e2 epitope (residues 120–

(A)

deletion   W28*   Q182*

Pre-C (29 aa) | Core (183 aa)
1814   1901

HBV DNA

Enh I
1060

DR2   SP1 & HNF3 binding   DR1

CP | URR | BCP
1613   1742

Enh II   31 nt duplication
1685   1775

deletion

XP | X (154 aa)
1230 1373   1835

              HNF1
BCP sequence   ATTAGGTTAAAGGTCTTT ← wild-type

    1753C   1762T   1766T
                1764A

aa in X mutant   I127T   K130M, V131I

(B)

	Pre-C (aa 1 -29)	Core (aa 1-183)
Occult group	C7G, L8V (3), P15S (3), P15L, V17F, G25R, W26R, W28*(7)	S12Y, V13A/L/M/Q, E14Q, L16I [S21A/Q, S26N (3), S26T, I27V (4), I27T], S35N, Y38H, E46D, {C48V (3), S49T (3), P50A (3), L55I (3), L55F, A58G, A58D, I59S, L60V (4)}, [S74G (3), E77D (2), E77Q, P79Q, A80T, E83D, L84A (3), S87G, S87R (2)], V91S, L95I/M, R98S, L100I (2), W102G, E113D/N/Q, L116I, [P130I, P130T (3), A131G/P, P135S], T147A, V149I (4), V149L, R151Q (4), R151G, G153S, S155T, P156S (2), Q169K, Q177K (3), E180G (2), Q182G (2), Q182*
Overt group	C12G, P15S, W28*	P5H (3), [S21A], S35T (7), Y38H, E46G, {P50A (5), H51Q, L60V (3)}, [A131P (4), P135Q], G153S, S176T, E180Q (2), Q182H (5), Q182*

Figure 3. Schematic diagrams showing types and positions of the mutations affecting the overlapping HBV key regulatory regions. (A) The overlapping regions of HBV genome contain core promoter (CP), basal core promoter (BCP), enhancer II (Enh II), DR1 (nt 1824–1834) and DR2 (nt 1590–1600) and coding regions of preC/C and X proteins. Clinically important mutations in BCP that results in the alteration of HNF1 binding site and the corresponding amino acid changes in the overlapping X gene. (aa: amino acid, nt: nucleotide, DR: direct repeat). Distinct mutations in TF binding sites, premature termination of protein synthesis, nucleotide duplication and deletion were indicated. Percentage of X mutants in overlapping regions in occult and control cases were listed. (B) Amino acid changes in the pre-C and core genes. (Position with more than one aa changes was bracketed).
doi:10.1371/journal.pone.0099028.g003

140) in OBI cases were also higher than controls [5/21 (23.8%) vs. 2/21 (9.52%), p<0.041].

Mutational Analysis on the X Coding Region

The total number of amino acid substitutions in the X ORF was significantly higher in subjects with OBI than controls [83/154 (53.9%) vs. 56/154 (36.4%), p<0.005]. A number of amino acid mutations were solely detected in subjects with OBI, including G22S (13.6%), P38S (13.6%), A102T (13.6%) and K118T (13.6%). There was a trend of a higher frequency of H86R mutation detected in OBI than in control cases [12/22 (54.5%) vs. 2/11 (9.1%), p = 0.067].

Mutational Analysis on the P Coding Region

The total number of amino acid substitutions over the P ORF was significantly higher in OBI than control cases [197/843 (30.1%) vs. 63/843 (7.6%), p<0.0001] (Table S2). Deletions in the pre-S/S region which lead to deletions and truncated proteins in the overlapping P coding region were found in 3/22 (13.6%) OBI cases. 3 out of 6 (50%) mutations detected in "α" determinant region of HBsAg also caused mutation in overlapping rt region, these included sI126S to rtD134E, sQ129N to rtS137Q and sT131N to rtN139K (Figure 2B). Nucleotide T128A mutation in the finger domain of the rt was uniquely detected in 3/22 (13.6%) OBI cases. It is speculated that rtT128A mutation might result in defective replication activity [35].

Distinct Mutations in the Regulatory Regions

The mutation frequency within the HBV key regulatory regions was comparable between the OBI group and control group (Table 1). However, a significant difference in the nucleotide diversity was noticed over the pre-S2/S promoter when comparing the OBI with the control groups with genotype C (p = 0.037; Table 2). Point mutations were the most noticeable changes that
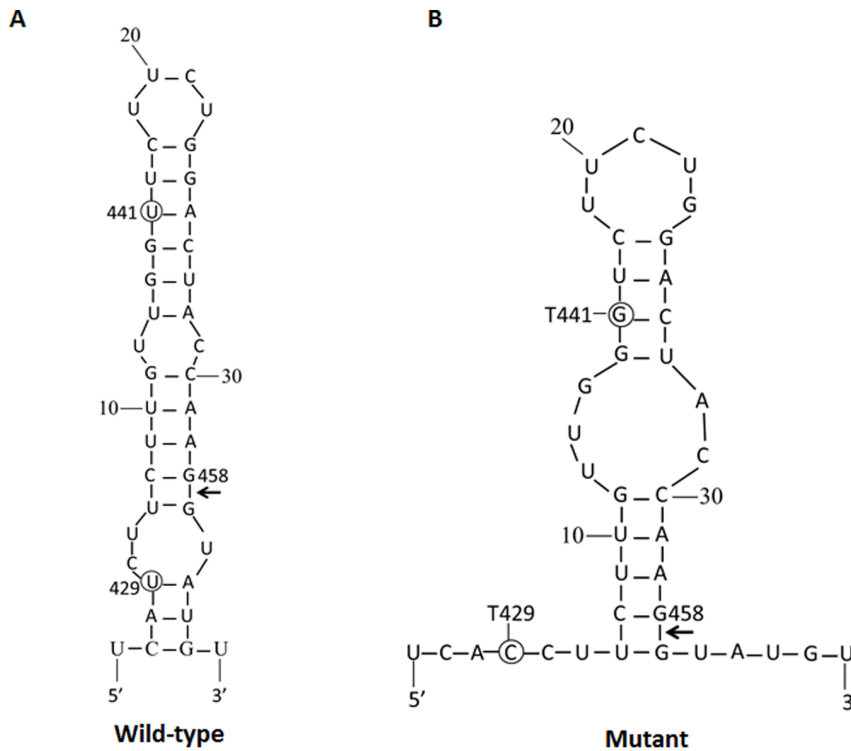
Figure 4. Prediction of RNA secondary structure at the 5′ splice site of occult hepatitis B virus infection (OBI) and control HBV strain. (A) RNA structure around the 5′ splice site at position 458 in control wild-type HBV strain (B) Predicted RNA structure of the OBI case with nucleotide T429C and T441G mutations. (Mutations are encircled and an arrow indicates the splice site).
doi:10.1371/journal.pone.0099028.g004

scattered over these key regulatory elements. Most of them result in interference with liver specific transcription factor (TF) binding sites and were more frequently detected in OBI cases than controls (Figure 2A, Figure 3A and Table S3). Large deletions at nt position 3127–55, 728–1255 and 1754–1771 were found in 3/22 (13.6%) OBI cases. These deletions led to loss of pre-S2/S and X promoters and disruption of BCP promoters, respectively. It is noteworthy that unique nucleotide duplications at the promoter regions were observed in 2/22 (9.1%) OBI cases. These unique nucleotide duplications included a 21 nt duplication (nt 3107–3127) in the Sp1 binding sites within the pre-S2/S promoter and a 31 nt duplication (nt 1644–1674) that interrupted the essential element box β in enhancer II (Table S3) [36].

Recent studies suggested that pre-S2/S mRNA splicing is essential for HBsAg expression [20,22]. This RNA splicing event is controlled by a 5′ splice donor site (nt 426–464) and the post-transcriptional element (PRE) that contains the 3′ splice acceptor site. Nucleotide G458A mutation was found to influence HBV RNA splicing and HBsAg production [20]. No G458A mutation was detected in this study. One OBI case with both T429C and T441G mutations in the 5′ splice donor site, which were predicted to affect mRNA secondary structure by MFOLD (Figure 4) [20]. The influence of RNA secondary structure on the activity of a 5′splice site of the S mRNA and optimal production of HBsAg needs further functional analysis.

Functional Analysis of the Distinct Nucleotide Alterations in Regulatory Elements

We next explored whether the distinct surface promoter and core promoter (CP) mutations identified in OBI cases caused a change in their level of transcription activity. Based on the findings of the mutations in the regulatory regions in OBI cases, 5 mutants and two wild-type controls were constructed. These included T2768A mutation located in Sp1 binding sites of pre-S1 (MUT1), 21 nt duplication between 3128–3148 on pre-S2/S (MUT2), C3015T mutation on preS2/S (MUT3) and two mutations G1677A (MUT4) and C1706T (MUT5) in CP. Luciferase activities of the plasmids containing these mutant promoters (MUT1–5) were compared to that of their corresponding wild-type promoter. The absolute luciferase activity of the control wild-type promoter was set to 100% and all the other mutant constructs were compared accordingly. As shown in Figure 5, mutation significantly affected the promoter activity in MUT2 (85%, $p < 0.05$), but not in MUT1 (92%, $p = 0.226$) and MUT3 (87%, $p = 0.156$). The promoter activity was significantly decreased in MUT4 (72%, $p < 0.05$) and MUT5 (40%, $p < 0.005$) with respective to the control CP.

Discussion

We performed mutational analysis of the RCA-amplified full-length HBV genomes from OBI and control cases. In order to exclude natural polymorphism and/or differences related to the geographical origin of the patients, it is imperative to have a robust comparison of sequences obtained from cases with the same genotype and origin. In this sense, we compared OBI sequences with both genotype-matched controls and reference sequences obtained from 49 genotype C (39 with sub-genotype C1 and 10 with sub-genotype C2) [26] and 5 genotype B2 CHB patients from Hong Kong [30].

Our full-length HBV sequence analysis revealed that HBV nucleotide diversity in OBI cases was significantly higher than control cases, indicating that more sequence variations were
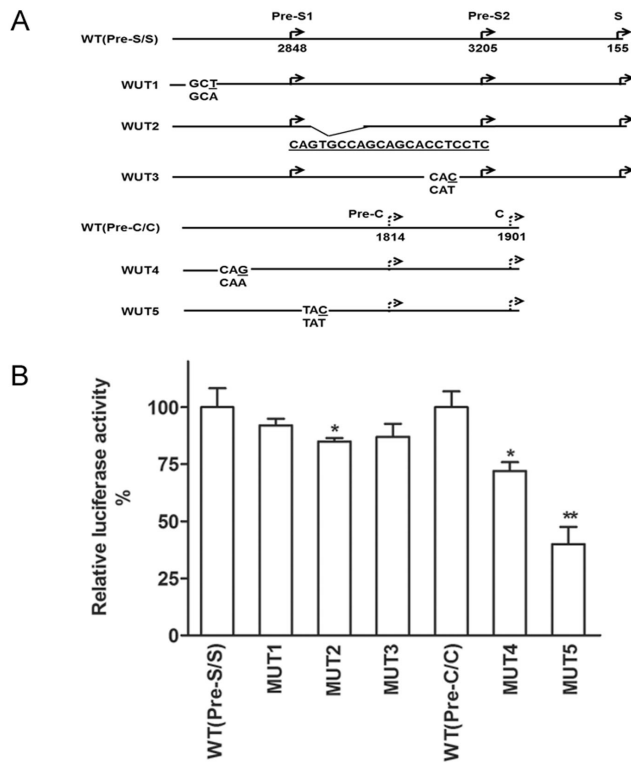
Figure 5. The relative luciferase activities of the constructs were determined in Huh-7 cells. (A) The HBV sequences showing type and position of the mutations in the constructs. (B) The relative luciferase activity in the mutant constructs. Error bars indicate the standard deviations. The luciferase activity of the wild-type (WT) was set to 100% and the changes with the different mutants were calculated accordingly.

doi:10.1371/journal.pone.0099028.g005

observed in OBI. Despite the strict criteria, a variety of mutations have been detected which are scattered throughout the viral genome in both occult and overt cases. Many of the mutations identified were in the surface proteins (Figure 2), which affect HBsAg detection assays, immune response recognition, HBV infectivity and virion morphogenesis [15,20,23]. However, we did not identify any single prevailing mutation or genetic signature which was associated with OBI. In this study, a mixture of mutation patterns like point mutations, deletions and nucleotide duplications were identified in OBI cases. This finding of multiple mutations, rather than prevailing mutations or genetic signatures, in OBI cases, is consistent with other reports [15,16,20,21,23,37]. This is further confirmed by our mutation frequency analysis that the amino acid mutation frequency within the individual genes and regulatory regions in the OBI cases was higher than that in the control cases. Taken together, it is likely that OBI is attributed to an increased accumulation of mutations in multiple regions, including key regulatory and coding regions, which in turn disrupts HBV replication and gene expression.

Despite the lack of prevailing OBI-associated mutations, this study identified several unique mutations within the HBV regulatory regions that may affect HBV replication. Many of these mutations identified in the key regulatory regions reside in the TF binding sites, which may disrupt TF bindings and hence affect promoter functions (Figure 2A and 3A). The reduced promoter activity was demonstrated in vitro by our luciferase assay, which showed that core promoter activity was reduced with

G1677A mutation in TF HNF3 binding sites and with C1706T mutation in the enhancer II region. We have also identified 2 OBI cases with duplications of HBV regulatory sequence. To our knowledge, this is the first description of nucleotide duplication in regulatory elements in OBI. This may explain the significant decrease in the pre-S2/S promoter activity caused by a 21 nt duplication between 3128–3148 by the luciferase assay.

This study also identified mutations in the "α" determinant region of the preS/S ORF in OBI subjects. Many mutations identified in the pre-S/S ORF, especially within the "α" determinant, have been demonstrated to affect the antigenicity and production of HBsAg, and thus possibly contribute to OBI [20,23,38,39]. This study identified an OBI case with the sG145A mutation, a vaccine escape variant which has been associated with OBI [33]. In addition, point mutations leading to the abolishment of pre-S2/S start codon and early stop signals, and sequence deletions resulting in truncated surface protein synthesis were also found in OBI cases. Xu et al. first reported that expression of a pre-S1 mutant would result in intracellular retention of HBsAg and that an appropriate balance of the various HBV surface proteins was important for functional HBsAg production and secretion [40,41]. Therefore, it is possible that the novel mutation patterns identified in the present study may disrupt surface protein synthesis and secretion. Of note, the change from sW182 to a stop codon (W182*) was identified in an occult case, which was recently suggested to be associated with OBI [42].

Due to the overlapping arrangement of HBV ORFs, mutations in the pre-S/S region could lead to amino acid changes in the overlapping P gene, especially in reverse transcriptase (rt). In particular, three immune escape mutations identified in "α" determinant region in OBI cases also caused mutations in the overlapping rt functional domains, including sI126S → rtD134E, sQ129N → rtS137Q and sT131N → rtN139K. The effect of these rt mutations on HBV replication level remains to be proven by further in vitro analysis.

Mutations associated with post-transcriptional regulation of HBsAg production were detected in OBI cases. The mechanism for the newly identified post-transcriptional regulation of HBsAg expression is still not yet fully elucidated. We have identified one occult case with T429C and T441G mutation at the 5′ donor splice site, which may affect the preS/S RNA secondary structure. As inhibition of HBsAg production could result from the mutation occurred in the pre-S2/S splicing sites [20,22,43], further investigation in larger OBI samples are needed.

In conclusion, our results indicated that subjects with OBI had a higher genetic diversity and higher mutation frequency than control subjects. The mutations identified in OBI cases included point mutations and sequence deletions, which might cause premature gene termination, defective HBsAg synthesis, decreased HBV promoter activity, or reduced HBV replication. All these mutations constrain viral replication capacity, and may collectively contribute to the OBI entity among the HBsAg-negative subjects. It is likely that multiple mechanisms are involved, and a combination of mutations, rather than any prevailing genetic signatures, is responsible for the OBI status.

Supporting Information

Figure S1 Phylogenetic analysis using the entire HBV genomes amplified from the 22 occult and 11 control HBV infected subjects. The filled circles represent the occult subjects, and the empty circles represent control subjects. Phylogenetic comparison was done by neighbor-joining algorithm based on Kimura two-parameter distance estimation. Bootstrap

values more than 75% are indicated on the major nodes. The scale of the evolutionary distances is shown at the bottom (scale bar). References sequences retrieved from GenBank are indicated by their accession numbers.
(DOCX)

Table S1 HBV-specific primers for rolling circle amplification (RCA), PCR amplification of full-length HBV genomes, sequencing and regulatory elements analysis.
(DOCX)

Table S2 Amino acid changes in HBV coding genes from subjects with occult and overt HBV infections.
(DOCX)

Table S3 Nucleotide changes in key regulatory regions of the hepatitis B virus (HBV) genome.
(DOCX)

Author Contributions

References

1. Moolla N, Kew M, Arbuthnot P (2002) Regulatory elements of hepatitis B virus transcription. J Viral Hepat 9: 323–331.
2. Tiollais P, Pourcel C, Dejean A (1985) The hepatitis B virus. Nature 317: 489–495.
3. Yuen MF, Wong DK, Fung J, Ip P, But D, et al. (2008) HBsAg Seroclearance in chronic hepatitis B in Asian patients: replicative level and risk of hepatocellular carcinoma. Gastroenterology 135: 1192–1199.
4. Da Silva LC, Madruga CL, Carrilho FJ, Pinho JR, Saez-Alquezar A, et al. (1996) Spontaneous hepatitis B surface antigen clearance in a long-term follow-up study of patients with chronic type B hepatitis. Lack of correlation with hepatitis C and D virus superinfection. J Gastroenterol 31: 696–701.
5. Liaw YF, Sheen IS, Chen TJ, Chu CM, Pao CC (1991) Incidence, determinants and significance of delayed clearance of serum HBsAg in chronic hepatitis B virus infection: a prospective study. Hepatology 13: 627–631.
6. Alward WL, McMahon BJ, Hall DB, Heyward WL, Francis DP, et al. (1985) The long-term serological course of asymptomatic hepatitis B virus carriers and the development of primary hepatocellular carcinoma. J Infect Dis 151: 604–609.
7. Lai CL, Shouval D, Lok AS, Chang TT, Cheinquer H, et al. (2006) Entecavir versus lamivudine for patients with HBeAg-negative chronic hepatitis B. N Engl J Med 354: 1011–1020.
8. Fattovich G, Giustina G, Sanchez-Tapias J, Quero C, Mas A, et al. (1998) Delayed clearance of serum HBsAg in compensated cirrhosis B: relation to interferon alpha therapy and disease prognosis. European Concerted Action on Viral Hepatitis (EUROHEP). Am J Gastroenterol 93: 896–900.
9. Allain JP (2004) Occult hepatitis B virus infection: implications in transfusion. Vox Sang 86: 83–91.
10. Raimondo G, Allain JP, Brunetto MR, Buendia MA, Chen DS, et al. (2008) Statements from the Taormina expert meeting on occult hepatitis B virus infection. J Hepatol 49: 652–657.
11. Raimondo G, Caccamo G, Filomia R, Pollicino T (2013) Occult HBV infection. Semin Immunopathol 35: 39–52.
12. Huo TI, Wu JC, Lee PC, Chau GY, Lui WY, et al. (1998) Sero-clearance of hepatitis B surface antigen in chronic carriers does not necessarily imply a good prognosis. Hepatology 28: 231–236.
13. Yuen MF, Wong DK, Yuan HJ, Sum SM, Lai CL (2004) HBsAg seroclearance in Chinese patients receiving lamivudine therapy for chronic hepatitis B virus infection. J Clin Microbiol 42: 4882–4884.
14. De Mitri MS, Cassini R, Bernardi M (2010) Hepatitis B virus-related hepatocarcinogenesis: molecular oncogenic potential of clear or occult infections. Eur J Cancer 46: 2178–2186.
15. Hou J, Wang Z, Cheng J, Lin Y, Lau GK, et al. (2001) Prevalence of naturally occurring surface gene variants of hepatitis B virus in nonimmunized surface antigen-negative Chinese carriers. Hepatology 34: 1027–1034.
16. Fang Y, Teng X, Xu WZ, Li D, Zhao HW, et al. (2009) Molecular characterization and functional analysis of occult hepatitis B virus infection in Chinese patients infected with genotype C. J Med Virol 81: 826–835.
17. Pollicino T, Amaddeo G, Restuccia A, Raffa G, Alibrandi A, et al. (2012) Impact of hepatitis B virus (HBV) preS/S genomic variability on HBV surface antigen and HBV DNA serum levels. Hepatology 56: 434–443.
18. Chaudhuri V, Tayal R, Nayak B, Acharya SK, Panda SK (2004) Occult hepatitis B virus infection in chronic liver disease: full-length genome and analysis of mutant surface promoter. Gastroenterology 127: 1356–1371.
19. Bock CT, Tillmann HL, Manns MP, Trautwein C (1999) The pre-S region determines the intracellular localization and appearance of hepatitis B virus. Hepatology 30: 517–525.
20. Candotti D, Lin CK, Belkhiri D, Sakuldamrongpanich T, Biswas S, et al. (2012) Occult hepatitis B infection in blood donors from South East Asia: molecular characterisation and potential mechanisms of occurrence. Gut 61: 1744–1753.
21. Allain JP, Belkhiri D, Vermeulen M, Crookes R, Cable R, et al. (2009) Characterization of occult hepatitis B virus strains in South African blood donors. Hepatology 49: 1868–1876.
22. Hass M, Hannoun C, Kalinina T, Sommer G, Manegold C, et al. (2005) Functional analysis of hepatitis B virus reactivating in hepatitis B surface antigen-negative individuals. Hepatology 42: 93–103.
23. Vivekanandan P, Kannangai R, Ray SC, Thomas DL, Torbenson M (2008) Comprehensive genetic and epigenetic analysis of occult hepatitis B from liver tissue samples. Clin Infect Dis 46: 1227–1236.
24. Margeridon S, Carrouee-Durantel S, Chemin I, Barraud L, Zoulim F, et al. (2008) Rolling circle amplification, a powerful tool for genetic and functional studies of complete hepatitis B virus genomes from low-level infections and for directly probing covalently closed circular DNA. Antimicrob Agents Chemother 52: 3068–3073.
25. Gunther S, Sommer G, Von Breunig F, Iwanska A, Kalinina T, et al. (1998) Amplification of full-length hepatitis B virus genomes from samples from patients with low levels of viremia: frequency and functional consequences of PCR-introduced mutations. J Clin Microbiol 36: 531–538.
26. Chan HL, Tsui SK, Tse CH, Ng EY, Au TC, et al. (2005) Epidemiological and virological characteristics of 2 subgroups of hepatitis B virus genotype C. J Infect Dis 191: 2022–2032.
27. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, et al. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28: 2731–2739.
28. Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, et al. (1999) Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. J Virol 73: 152–160.
29. Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 31: 3406–3415.
30. Sugauchi F, Orito E, Ichida T, Kato H, Sakugawa H, et al. (2002) Hepatitis B virus of genotype B with or without recombination with genotype C over the precore region plus the core gene. J Virol 76: 5985–5992.
31. Carman WF (1997) The clinical significance of surface antigen variants of hepatitis B virus. J Viral Hepat 4 Suppl 1: 11–20.
32. Coleman PF, Chen YC, Mushahwar IK (1999) Immunoassay detection of hepatitis B surface antigen mutants. J Med Virol 59: 19–24.
33. Martin CM, Welge JA, Rouster SD, Shata MT, Sherman KE, et al. (2012) Mutations associated with occult hepatitis B virus infection result in decreased surface antigen expression in vitro. J Viral Hepat 19: 716–723.
34. Pollicino T, Raffa G, Costantino L, Lisa A, Campello C, et al. (2007) Molecular and functional analysis of occult hepatitis B virus isolates from patients with hepatocellular carcinoma. Hepatology 45: 277–285.
35. Hunt CM, McGill JM, Allen MI, Condreay LD (2000) Clinical relevance of hepatitis B viral mutations. Hepatology 31: 1037–1044.
36. Kramvis A, Kew MC (1999) The core promoter of hepatitis B virus. J Viral Hepat 6: 415–427.
37. Lee WM (1997) Hepatitis B virus infection. N Engl J Med 337: 1733–1745.
38. Weinberger KM, Bauer T, Bohm S, Jilg W (2000) High genetic variability of the group-specific a-determinant of hepatitis B virus surface antigen (HBsAg) and the corresponding fragment of the viral polymerase in chronic virus carriers lacking detectable HBsAg in serum. J Gen Virol 81: 1165–1174.
39. Yamamoto K, Horikita M, Tsuda F, Itoh K, Akahane Y, et al. (1994) Naturally occurring escape mutants of hepatitis B virus with various mutations in the S gene in carriers seropositive for antibody to hepatitis B surface antigen. J Virol 68: 2671–2676.
40. Xu Z, Yen TS (1996) Intracellular retention of surface protein by a hepatitis B virus mutant that releases virion particles. J Virol 70: 133–140.
41. Xu Z, Jensen G, Yen TS (1997) Activation of hepatitis B virus S promoter by the viral large surface protein via induction of stress in the endoplasmic reticulum. J Virol 71: 7387–7392.
42. Lee SA, Kim K, Kim H, Kim BJ (2012) Nucleotide change of codon 182 in the surface gene of hepatitis B virus genotype C leading to truncated surface protein is associated with progression of liver diseases. J Hepatol 56: 63–69.
43. Sommer G, Heise T (2008) Posttranscriptional control of HBV gene expression. Front Biosci 13: 5533–5547.