Published in final edited form as:

Virus Res. 2017 May 02; 235: 86–95. doi:10.1016/j.virusres.2017.03.021.

Sequence analysis and functional characterization of full-length hepatitis B virus genomes from Korean cirrhotic patients with or without liver cancer

Huailiang Zhou^{#a}, Dina Gewaily^{#b}, Sang Hoon Ahn^c, Carina Preskill^b, Yongxiang Wang^a, Li Zong^a, Jing Zhang^a, Kwang-Hyub Han^c, Jack Wands^b, Jisu Li^b, and Shuping Tong^{a,b,*}

^aKey Lab of Medical Molecular Virology School of Basic Medical Sciences Fudan University Shanghai, China

^bLiver Research Center, Rhode Island Hospital, Warren Alpert School of Medicine, Brown University, Providence, RI, USA 1661753336@qq.com

^cDepartment of Internal Medicine, Institute of Gastroenterology, Yonsei University College of Medicine, Seoul, Republic of Korea

[#] These authors contributed equally to this work.

Abstract

This study aimed to identify and characterize mutations in the hepatitis B virus (HBV) genome associated with advanced liver diseases. The 3.2-kb HBV genome of the C2 subgenotype was amplified from sera of 18 cirrhotic Korean patients with (10) or without (8) hepatocellular carcinoma (HCC), and two clones per patient were characterized by transient transfection experiments in human hepatoma cells. While A1762T/G1764A core promoter mutations were highly prevalent in both groups, the G1896A precore mutation to abolish hepatitis B e antigen (HBeAg) expression was more common in HCC clones (55% vs. 20%). High replication capacity was mostly found in HCC clones and associated with core promoter mutations, whereas more non-HCC clones harbored a nonfunctional core gene (34% vs. 8%). Large in-frame deletions in the preS region were found in 60% of HCC clones and 38% of non-HCC clones. They removed the first 11 residues of large envelope protein or impaired small envelope protein expression, or deleted a neutralizing epitope in the preS2 domain. Additional point mutations prevented middle envelope protein expression, or caused nonsense mutations in the preS or S region to truncate large and/or small envelope protein. Consequently, many clones were unable to express or secrete hepatitis B surface antigen (HBsAg). In conclusion, mutations associated with the advanced stage of chronic HBV infection are complex and diverse. Host immune pressure most likely selected for mutations in the HBV genome to abolish or reduce HBeAg or HBsAg production, to enhance genome replication, or to escape neutralizing antibodies. Some of these mutations may contribute to liver cirrhosis or HCC development.

*Corresponding author at: Liver Research Center, Rhode Island Hospital, 55 Claverick Street, Providence, RI 02903, USA.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.virusres.2017.03.021>.

Keywords

Hepatitis B virus; Liver cancer; Genome replication; Core promoter mutations; PreS deletions; Hepatitis B surface antigen

1. Introduction

Chronic hepatitis B virus (HBV) infection may lead to the development of liver cirrhosis and hepatocellular carcinoma (HCC). Majority of the 240 million chronic HBV carriers live in East Asia, where the perinatal mode of HBV transmission greatly promotes chronicity of infection. The HBV genome of 3.2 kb contains four genes: precore/core, preS1/preS2/S (envelope), polymerase (P), and X. The core protein is assembled into core particle essential for genome replication, while the precore/core protein is processed and secreted as hepatitis B e antigen (HBeAg). Similarly, alternative translation initiation from ATG codons in the envelope gene generates preS1/preS2/S (large or L), preS2/S (middle or M), and S (small) envelope proteins. At the mRNA level, HBeAg is translated from the 3.5-kb precore RNA (pcRNA), while core protein and DNA polymerase are expressed from the slightly shorter pregenomic RNA (pgRNA). The envelope proteins are translated from the 2.4-kb (L) and 2.1-kb (M and S) subgenomic RNAs. Transcription of the 3.5-kb, 2.4-kb, and 2.1-kb RNAs is driven by the core, SPI, and SPII promoters, respectively. HBV genome replication is driven solely by the pgRNA, which, together with P protein, is co-packaged into newly assembled core particle. The P protein then catalyzes minus strand DNA synthesis, pgRNA degradation, and plus strand DNA synthesis. Thus, HBV replication capacity is largely controlled at the transcriptional level.

Interaction of core particles with envelope proteins anchored on ER membrane leads to virion formation and release. Of the three envelope proteins, L is involved in interaction with core particles, while S drives particle release. In the absence of L protein, S protein is secreted alone as noninfectious subviral particles. L protein can suppress subviral particle secretion depending on the L/S protein ratio. Since S protein expression far exceeds that of L protein, during natural HBV infection subviral particles outnumber virions by a factor of 1000:1 or so. The envelope proteins present on virions and subviral particles can be detected by anti-S antibodies as hepatitis B surface antigen (HBsAg), which provides a sensitive marker of HBV infection.

The immune tolerance phase of chronic HBV infection is characterized by high HBV DNA titer and presence of both HBeAg and HBsAg in the blood. HBeAg seroconversion (loss of HBeAg followed by rise of corresponding antibody) during the immune clearance phase often leads to marked drop in HBV DNA titer. Subsequent HBsAg seroconversion indicates the resolution of infection. Being a soluble protein rather than a component of HBV virions, HBeAg is not essential for HBV life cycle but rather helps induction of immune tolerance. The G1896A nonsense mutation in the precore region, which blocks HBeAg expression at the translational level, is frequently selected at the HBeAg- negative phase of chronic HBV infection (Tong et al., 2013). Alternatively, HBeAg expression can be down regulated by mutations in the core promoter such as A1762T/G1764A, which reduce transcription of the

pcRNA. Core promoter mutations can also up regulate genome replication through increased transcription of pgRNA (Baumert et al., 1996; Parekh et al., 2003; Tsai et al., 2009). Finally, the rising anti-HBs antibodies can select for amino acid substitutions in the antigenic epitopes in the S protein such as G145R. Such mutations greatly diminish affinity of the S protein for anti-HBs antibodies elicited by wild-type S protein, thus could lead to immune escape and vaccine failure (Carman et al., 1990).

Molecular epidemiological studies implicated high viral load and HBeAg persistence as risk factors for HCC development. As for the genetic variability in the HBV genome, core promoter mutations (Baptista et al., 1999; Kao et al., 2003) and deletions in the preS region (Fang et al., 2008; Chen et al., 2008) are more prevalent in HCC than non-HCC samples. HBV isolates worldwide can be grouped into ten genotypes according to sequence divergence in the entire genome (Kramvis, 2014; Norder et al., 2004; Tong and Revill, 2016). Genotypes B and C prevail in East Asia, an area of high HCC incidence. Compared with genotype B infection, infection with genotype C is associated with prolonged replicative phase, delayed HBeAg seroconversion, higher prevalence of both core promoter mutations (Orito et al., 2001; Yuen et al., 2003) and preS deletions (Chen et al., 2006; Chen et al., 2008), and consequently increased risk for HCC development. However, epidemiological studies on HBV-HCC relationship are often complicated by host genetic background and environmental factors. More importantly, little functional characterization of full-length HBV genomes isolated from HCC patients has been performed, largely due to many steps required to convert the full-length genome amplified from clinical samples into replication competent tandem dimer (Parekh et al., 2003). The present study focused on Korean chronic HBV carriers who are nearly invariably infected with the C2 subgenotype (Ahn et al., 2010; Kim et al., 2009). Full-length HBV genome was amplified from 10 HBeAg-positive HCC patients with underlying liver cirrhosis, as well as 9 age- and gender-matched cirrhotic patients without HCC. A total of 36 clones were transiently transfected to human hepatoma cell lines to characterize their HBeAg and HBsAg production, as well as genome replication.

2. Materials and methods

2.1. Serum samples, amplification of the entire HBV genome, and conversion to a replication competent form

Serum samples were collected with consent from 10 patients diagnosed with liver cirrhosis and HCC, as well as 10 cirrhotic patients without HCC who were nevertheless paired with the HCC patients in terms of gender and age. The study was approved by the Ethics Committee of Severance Hospital, Yonsei University. HBsAg and HBeAg titers from these samples were measured using Auszyme Monoclonal kit (Abbott Laboratories) and ETI-EBK plus kit (Diasorin), respectively. DNA was extracted from 200 μ l of serum samples using QiAmp DNA Blood Mini Kit (Qiagen), and the full-length HBV genome was amplified by 40 cycles of polymerase chain reaction (PCR) using High Fidelity PCR system (Roche) and sense primer 5'-CCGGAAAGCTTATGCTCTTCTTTTTCACCTCTGCCTAATCATC-3' (HindIII site underlined) and antisense primer 5'-CCGGAGAGCTCATGCTCTTCAAAAAGTTGCATGGTGCTGGTG-3' (SacI site

underlined) (Parekh et al., 2003). The PCR product was cloned to the HindIII and SacI sites of pUC18 vector, and two clones per serum sample were converted to replication competent dimer construct through several steps of enzymatic manipulation (Parekh et al., 2003). The resultant SphI dimer has two copies of the HBV genome joined in tandem via the unique SphI site on the viral genome and linked to pUC18 through the same site. The SphI dimer could generate the 3.5-kb terminally redundant pcRNA and pgRNA required for HBeAg expression and genome replication, respectively. The full-length sequences of these 36 HBV clones were determined and deposited to GenBank under the accession numbers KY363252-KY363287.

2.2. Subgenomic expression constructs and site-directed mutants

A 0.7mer construct for expression of the three envelope proteins under their own promoter and enhancer elements has already been described (Garcia et al., 2009). For the present study, the 1.3-kb SacI/ HindIII fragment covering the envelope gene was replaced with the cognate fragment (positions 2715–835) of clone 2.6. The primers for amplification of the envelope gene were sense 5'-CATGGAGCTCGCA- GTTAATCATTACTTCAAAAC (SacI site underlined)-3' and antisense 5'- CATGAAGCTTAAATGTATACCCAAAGA-3' (HindIII site underlined). Various deletions or point mutations were introduced to this 0.7mer HBsAg expression construct by overlap extension PCR. Site-directed mutagenesis was also used to mutate the precore and/or core promoter sequence from SphI dimer of clone 4.2 and clone 19.3 back into wild type sequence. To express HBeAg under the CMV promoter, the precore/core gene (positions 1787–2482) of clone 16.3 was amplified by PCR using primer pair 5'-CATGGAGCTCGCATAAATTGGTCTGTTC- ACCAGCACCATGC-3' (SacI site underlined) and 5'-CATGAAGCTTAA- GTTTCCACCTTATGAGTCCAAG-3' (HindIII site underlined). The PCR product was cloned to the SacI/HindIII sites of pcDNA3.1zeo (—) vector (Ito et al., 2009). Various mutations and deletions in the precore region or core gene were introduced into this construct to test their impact on HBeAg expression.

2.3. Transfection of human hepatoma cells and analysis of viral DNA replication and protein expression

Huh7 cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (Gibco), while HepG2 cells were cultured in Minimum Eagle's medium (MEM). Cells seeded overnight in 6-well plates were transfected with 2 µg of HBV DNA using either transIT-LT1 reagent (Mirus) or Lipofectamine 2000 (invitrogen), and harvested 4 or 5 days post-transfection. HBsAg and HBeAg were measured from culture supernatant with or without dilution using ELISA kits (KHB, Shanghai, China). Cells were lysed in 80 µl of buffer containing 10 mM Hepes (pH7.5), 100 mM NaCl, 1 mM EDTA, and 1% NP40. Half of the cell lysate was used for extraction of replicative HBV DNA (Parekh et al., 2003; Garcia et al., 2009). Briefly, the samples were digested at 37 °C for 15 min with 5U of mung bean nuclease and 1U of DNase I in the presence of 10 mM CaCl₂ and 12 mM MgCl₂, so as to degrade transfected plasmid DNA. Core particles were precipitated with polyethylene glycol (PEG) solution (1.2 M NaCl, 60 mM EDTA, 30% sucrose, 26% PEG8000), resuspended in a solution containing 10 mM Tris (pH7.5), 6 mM MgCl₂, and 8 mM CaCl₂, and further treated with mung bean nuclease (3U) and DNase I (1U) at 37 °C for 15 min. Next, core particles were disrupted by addition of proteinase K buffer (25 mM

Tris, pH7.5, 10 mM EDTA, 100 mM NaCl, 0.5% SDS), and proteins were digested at 37 °C for 2 h with 0.5 mg/ml of proteinase K. Samples were extracted with Tris-saturated phenol, and nucleic acids were precipitated with ethanol. After gel electrophoresis and Southern transfer, the blots were hybridized with HBV DNA probe of genotype C generated by random priming from a 3.1-kb pure HBV DNA fragment.

For Western blot, proteins were separated in SDS-12% polyacrylamide gel (PAGE) using one fourth of the cell lysate. After transfer, the blot was incubated overnight at 4 °C with a polyclonal rabbit anti-HBs antibody (Novus) diluted 1:4000 in 5% milk/TBST. The blot was washed and incubated at room temperature for 1hr with a 1:10,000 dilution of goat anti-rabbit antibodies conjugated with horseradish peroxidase (HRP, Amersham). Signals were revealed by enhanced chemiluminescence (ECL). The blot was treated with stripping buffer, washed with TBST, and reprobed with polyclonal mouse anti-p-actin antibody followed by goat anti-mouse antibody conjugated to HRP. L protein was detected by a custom made rabbit polyclonal antibody against residues 12–46 in the preS1 domain (Li et al., 2016), which was used at 1:4000 dilution.

3. Results

3.1. Overview

We performed PCR amplification of the full-length HBV genome from 20 Korean cirrhotic patients with or without HCC. Following molecular cloning, partial DNA sequencing was performed to check for sequence integrity in the core gene. Two clones per patient with a functional core gene were converted to SphI dimer, which were fully sequenced to identify mutations or deletions/insertions at the nucleotide and amino acid levels. Moreover, the full-length clones were transiently transfected to human hepatoma cell line Huh7 or HepG2 to characterize their properties of genome replication, envelope protein expression, as well as HBeAg and HBsAg secretion. To elucidate why many clones failed to secrete HBeAg or HBsAg, subgenomic expression constructs for envelope proteins or HBeAg were generated and site-directed mutagenesis was performed.

3.2. HCC clones often maintained a functional core gene but had higher prevalence of mutations affecting the last four residues of core protein

Serum samples were obtained from 10 HCC patients and 10 non-HCC patients. Both groups of patients had liver cirrhosis, with the non-HCC patients matched one-to-one with the HCC patients in terms of gender, sex, and age (Table 1). PCR amplification of the full-length HBV genome was successful for all samples except for #9 from the non-HCC group. At least two PCR clones per sample were sequenced with the P1 primer to cover the entire core gene of 183 codons, and additional clones were sequenced if one or both had a nonfunctional core gene. Of the 79 clones thus sequenced for the core gene (Table 2), 17 had a nonfunctional core gene due to mutated initiation codon (1), frameshift (2), large internal in-frame deletion (7), and nonsense mutations to greatly shorten the core protein (7). Such a nonfunctional core gene was more prevalent in non-HCC clones (14/41, or 34%) than in HCC clones (3/38, or 8%) (Fig. 1). Conversely, the Q182* nonsense mutation, which shortened core protein by just two residues without impairing its function in supporting

genome replication (Nassal, 1992), was present in 8 HCC clones but only 1 non-HCC clone. In addition, the E180G, S181P, and Q182K missense mutations were identified in 12 clones derived from 5 HCC patients but only 3 clones from 2 non-HCC patients (Table 2).

3.3. HCC clones showed high prevalence of G1896A mutation in association with low serum HBeAg titers

Sequencing using the P1 primer also allowed us to read DNA sequence at the 3' precore region, where the G1896A hot spot mutation to abolish HBeAg expression is located. This mutation was identified in 55% of the HCC clones analyzed (21/38), but only 20% of the non-HCC clones (8/41) (Fig. 1). In this regard the HBeAg titer was quite low for 6 of the 10 HCC samples (#1, #5, #11, #13, #17, #19) but only 2 non-HCC samples (#8, #18) (Table 1). Among samples with extremely low HBeAg titers, G1896A was present in all the clones from patient #1, #11, #13, #17, and most clones from #18. For #5, one clone harbored G1896A, while the other clone had deletion of codons 84–121. For patient #8, 3 of the 5 clones had a nonsense mutation at codon 36 in the core gene (Table 2), which should abolish both genome replication and HBeAg production. Therefore, the low HBeAg titers in most serum samples correlated with the high prevalence of G1896A mutation to abolish HBeAg expression.

3.4. Majority of HCC clones and half of the non-HCC clones available as SphI dimer were deficient in HBeAg production

Two clones per sample with a functional core gene were converted to SphI dimer. The only exception was clone 5.2, as we would like to test the replication impact of the large in-frame deletion in the core gene. Due to the negative PCR for serum #9 and defective core gene for all the clones from serum #10, only 16 clones from 8 non-HCC patients were analyzed together with 20 clones from 10 HCC patients. HBeAg was undetectable from culture supernatant of Huh7 cells transfected with the 18 clones harboring the G1896A mutation, even without sample dilution (Supplementary Fig. 1). For the remaining 18 clones, HBeAg from undiluted culture supernatant was negative or extremely low for clones 5.2, 12.1, 12.3, 14.3, and 15.5. Together, 75% (15/20) of HCC clones and 50% (8/16) of non-HCC clones were negative or nearly negative in HBeAg production (Fig. 1; Supplementary Table 1). To establish the structural basis, the entire precore/core gene of clone 16.3 with efficient HBeAg expression was inserted to pcDNA3.1 vector for HBeAg production under the CMV promoter. Deleting codons 84–121 in the core gene, as found in clone 5.2, caused HBeAg-negative phenotype (Supplementary Fig. 1B). On the other hand, introducing the S155P mutation (found in clones 12.1 and 12.3) failed to impair HBeAg expression.

3.5. Core promoter mutations were highly prevalent in both HCC and non-HCC clones

The HBeAg-negative phenotype of clones 12.1, 12.3, 14.3, and 15.5 could not be explained by nonsense or frameshift mutations in the precore or core gene. In this regard core promoter mutations could reduce HBeAg expression at the transcriptional level (Parekh et al., 2003; Tsai et al., 2009). The most common (hot spot) core promoter mutations are A1762T in conjunction with G1764A. Consistent with the high prevalence of core promoter mutations in genotype C, especially at the advanced stages of liver diseases (Orito et al., 2001; Yuen et al., 2003), 34 of the 36 clones (including all the 20 HCC-derived clones) harbored the

A1762T/G1764A double mutation (Fig. 1; Supplementary Table 1). Among these clones, 13 (9 from HCC patients) harbored additional T1753C mutation (Fig. 1), 2 had additional C1766T mutation. Two clones derived from HCC samples (3.5 and 4.9) had 24-nt insertion between 1662/1663 and 11-nt insertion between 1628/1629, respectively. Finally, the only two clones lacking A1762T/G1764A mutations (12.1 and 12.3 from a non-HCC patient) had nt 1727 and 1768–1775 deleted (Supplementary Table 1). These two clones happened to be HBeAg negative despite the lack of a G1896A nonsense mutation, although a causative role of core promoter deletions on negative HBeAg phenotype was not determined.

3.6. Insertion or deletion in the core promoter or precore region markedly changed the primary sequence of HBx protein

The entire core promoter and the 5' end of the precore region overlap with the X gene. Thus, the 24-nt insertion in clone 3.5 caused insertion of 8 residues between positions 93 and 94 in the 154-aa HBx protein (Table 3). The 11-nt insertion in clone 4.9 would have caused frameshift mutation, but this clone could not express HBx protein due to mutated ATG codon. Clones 2.10 and 11.2 had 1-nt deletion at position 1834 and 1821 in the precore region, respectively, which resulted in the C-terminal extension of HBx protein to variable lengths (Table 3). Finally, deletion of nt 1727 in clones 12.1 and 12.3 led to fusion between the N-terminal 118 residues of HBx protein and 10 residues of irrelevant sequence.

3.7. More HCC than non-HCC-derived clones possessed high replication capacity, which could be attributed to core promoter mutations

Transient transfection experiments in Huh7 and HepG2 cells revealed 11 clones to be defective or inefficient in genome replication: 5 from HCC patients (2.6, 2.10, 5.2, 15.5, 17.4) and 6 from non-HCC patients (7.7, 7.9, 14.2, 14.3, 18.1, 18.2) (Fig. 2A and B; Supplementary Table 1). The relative replication capacities for the remaining 25 clones were more or less consistent between Huh7 and HepG2 cells, with efficient replication achieved by clones 1.5, 1.6, 4.2, 4.9, 5.1, 6.1, 6.3, 11.1, 11.2, 13.2, 16.3, and 19.3. All but clones 6.1, 6.3, 16.3 were derived from HCC patients, indicating a higher percentage of HCC clones were high replicating (9/20, or 45%) than non-HCC clones (3/16, or 19%) (Fig. 1). All the high replicating clones harbored A1762T/G1764A mutations, and all but the three non-HCC clones harbored G1896A precore mutation. Core promoter mutations not only suppress transcription of pcRNA, but also augment pgRNA transcription, thus increasing genome replication (Baumert et al., 1996). Our previous studies found that T1753C and C1766T could enhance the replication promoting effect of A1762T/G1764A hot spot mutations in genotype A (Parekh et al., 2003; Tsai et al., 2009). In this regard 7 of the 12 highest replicating clones harbored either T1753C or C1766T. To dissect the relative contribution of precore vs. core promoter mutations on the high replication capacity, back mutations were introduced to clones 4.2 and 19.3. The T1762A/A1764G back mutations markedly reduced genome replication for clone 4.2, whereas the A1896G back mutation had limited effect (Fig. 2C). Reversing both core promoter and precore mutations had similar effect as reversing core promoter mutations alone. Similar results were obtained for clone 19.3, which had T1753C/A1762T/G1764A triple core promoter mutations (Fig. 2C).

3.8. HCC clones had high prevalence of in-frame deletions in the preS region

The preS region is subdivided into the preS1 region (nt 2848–3204 for genotype C) and preS2 region (nt 3205–154). Eighteen of the 36 clones harbored deletions in the preS region: 7 in preS1 (4.9, 7.9, 11.1, 11.2, 12.1, 12.3, 17.1), 7 in preS2 (2.10, 3.5, 3.6, 6.1, 13.1, 13.2, 15.3), and 4 in both regions (8.3, 8.4, 15.5, 19.3) (Fig. 3A; supplementary Table 2). Twelve of these 18 clones originated from HCC patients, while the remaining 6 clones came from 4 non-HCC patients. Therefore, the preS deletions were more prevalent in HCC-derived clones (60% vs. 38%; Fig. 1). The deletions ranged from 6 to 144 nt, with all being multiples of 3. As can be seen in Fig. 3A, the 5' end of the preS1 region and the 5'/central part of preS2 region are the deletional hot spots. Other deletions partially or completely overlap with the SPII promoter. In terms of protein expression, deletions at the 5' end of the preS1 region all eliminated the initiating ATG. Since there is another in-frame ATG codon 33 nucleotides downstream, the deletions did not abolish L protein expression but rather shortened L protein by 11 residues (Fig. 3B). On the other hand, 4 clones derived from two HCC patients (2.6, 5.1, 5.2) and one control patient (6.3) could not express M protein due to mutated preS2 ATG (Fig. 4).

3.9. PreS deletions removing the 3' end of the SPII promoter could cause HBsAg-negative phenotype

The 36 full-length HBV clones as SphI dimer were transfected to Huh7 cells, and HBsAg was measured from both culture supernatant and cell lysate by ELISA. Five clones each from the HCC group (4.9, 5.2, 15.5, 17.1, 19.3) and non-HCC group (7.9, 8.3, 8.4, 12.1, 14.3) failed to produce HBsAg in both cell lysate and culture supernatant (Fig. 1; Fig. 5B&C), even when 1:10 diluted cell lysate and undiluted culture supernatant were used (supplementary Table 2). Western blot analysis also failed to detect S protein in cell lysate (Fig. 5A). Two additional clones (1.5 and 12.3) had HBsAg in cell lysate but not in culture supernatant, even when assayed undiluted. Among the 10 clones lacking both intracellular and secreted HBsAg, all but clones 5.3 and 14.3 harbored preS deletions (Supplementary Table 2). In this regard the 2.1-kb RNA used for S protein expression is driven by the SPII promoter located in the preS1 region (Fig. 3A). A 0.7mer expression construct for the envelope proteins of clone 2.6 under endogenous HBV enhancers and promoters was generated, and impact of some of the preS deletions was evaluated by site-directed mutagenesis. S protein expression and HBsAg production were somewhat reduced by deleting nt 2890–2976 or 2973–3083, and abolished by deletion covering nt 3123–3164 and 3132–60 (Fig. 6). Therefore, deletions 3123–3164 and 3132–60 probably contribute to the HBsAg-negative phenotype of clones 17.1 and 19.3 (Table 4).

3.10. Premature nonsense mutations or missense mutations in the S region could also lead to HBsAg-negative phenotype

Of the 10 clones with HBsAg-negative phenotype, 8 harbored a nonsense mutation at codon 61, 69, 172, 182, 201, or 209 in the S region (Fig. 4), while another clone, 7.9, had a +1 frameshift mutation fusing the N-terminal 188 residues of S protein with the C-terminus of P protein. Through site-directed mutagenesis, we found that intracellular S protein and HBsAg were diminished by nonsense mutation at codon 209, 182, and especially 201, and abolished

by a nonsense mutation at codon 69 and 61 or by the frameshift mutation (Fig. 6). These findings could explain the HBsAg-negative phenotype of many clones (Table 4). Using undiluted culture supernatant, it was evident that a nonsense mutation at codon 209 completely abolished HBsAg secretion, whereas a stop codon at position 201 and 182 did not (data not shown). Some of the SphI dimers with a negative HBsAg phenotype harbored missense mutations in the S region of the envelope gene (Supplementary Table 2). In particular, clone 14.3 harbored a D99G missense mutation but no preS deletion or nonsense mutation in the envelope gene. The impact of several missense mutations on HBsAg production and secretion was analyzed by site-directed mutagenesis. The P120T, I126S, and R160G mutations did not affect intracellular HBsAg level, although R160G slightly reduced extracellular HBsAg (Fig. 6B). The G145R immune escape mutation caused low S protein level by Western blot. Remarkably, the D99G mutation completely abolished HBsAg for both cell lysate and culture supernatant (Fig. 6), which could explain the HBsAg negative phenotype of clone 14.3 (Table 4).

3.11. Nonsense mutations in the S or preS region could prevent the translation of full-length L protein

Full-length L protein could not be detected from lysate of Huh7 cells transfected with 6 clones from the non-HCC group (7.9, 8.3, 8.4, 12.1, 12.3, 14.3) and 6 clones from the HCC group (4.9, 5.2, 15.5, 17.1, 17.4, 19.3) (Fig. 5A). DNA sequencing identified mutations responsible for the L-negative phenotype for at least 9 clones (Table 4). These include lost initiation sites (19.3) and nonsense mutation at codon 121 or 72 in the preS region (8.3, 8.4, 12.1, 12.3) (Fig. 4). Four clones (4.9, 5.2, 7.9, 15.5) had truncation or frameshift inside the S domain, which abolished S protein expression as well. This may partly explain why clones unable to express S protein were often deficient in L protein expression as well.

4. Discussion

In East Asia where HBV genotypes B and C are perinatally transmitted, liver cirrhosis and especially HCC are the consequences of decades of virus replication and repeated cycles of immune-mediated destruction of infected hepatocytes. Thus, viral quasispecies detected at such a late stage of infection may have accumulated mutations in many parts of the viral genome, with some of the mutations directly or indirectly driving liver pathogenesis. In the present study, we sequenced full-length HBV genome from 20 HCC-derived clones and 16 non-HCC clones (Supplementary Table 1&2), and also sequenced 3' precore/entire core gene from 18 additional HCC-derived clones and 25 additional non-HCC clones (Table 2). The G1896A precore mutation to abolish HBeAg expression occurred more often in HCC clones (Fig. 1), although that might be partly attributed to the extremely low HBeAg titer found in many HCC samples (Table 1).

Internal in-frame deletions in the core gene were found primarily from non-HCC clones. These deletions covered residues 84–121, 88–128, 89–128, and 79–107 (Table 2). In this regard, a T cell epitope has been mapped to residues 85–100 (Milich et al., 1987), and codons 84–101 were hot spots for missense mutations in hepatitis B patients with liver diseases (Ehata et al., 1992). We found that clone 5.2 lacking residues 84–121 was

replication deficient (Fig. 2), and deleting this sequence from an HBeAg expression construct caused HBeAg-negative phenotype (Supplementary Fig. 1B). Thus, the large deletion apparently abolished core protein function. Nonsense mutations to truncate a large portion of the core protein were also more prevalent in non-HCC clones (Table 2). Even excluding clones with a defective core gene from dimer construction, subsequent transfection experiments of the 36 clones as SphI dimer revealed association of non-HCC clones with reduced replication capacity (Fig. 2). Among the 11 clones defective or inefficient in genome replication, 6 were from non-HCC patients. Since both clones from patient #2, #7, #14, and #18 were replication deficient, a role of PCR errors for these 8 clones could be largely ruled out. In this regard, 3 of the 4 patients (#7, #14, and #18) were from the non-HCC group. Among the 12 high replicating clones, only 3 were from the non-HCC group. The 9 high replicating HCC clones originated from 6 serum samples, with 7 clones from 5 serum samples containing low HBeAg titers (#1, #5, #11, #13, #19). Thus HBV genomes from HCC patients are more likely to maintain a functional core gene and to possess high replication capacity, especially those with low HBeAg titers. Site-directed mutagenesis using clones 4.2 and 19.3 confirmed a role of core promoter mutations (A1762T/G1764A in clone 4.2, and T1753C/A1762T/G1764A in clone 19.3) in the high replication capacity (Fig. 2C). This is consistent with previous reports based on other HBV genotypes (Baumert et al., 1996; Parekh et al., 2003).

Interestingly, HCC clones were more likely to possess mutations affecting the last 4 residues of core protein (Table 2). These include E180G, S181P, Q182K, and Q182* to remove the last 2 residues (Q182 and C183) of core protein. The functional significance of such mutations remains unclear. The C-terminal 10 residues of the core protein are dispensable for HBV replication (Nassal, 1992), and the high replication capacity of clones 4.2, 4.9, 13.2, and 5.1 suggests that the E180G, S181P, and Q182K mutations present in these clones do not markedly suppress genome replication.

A large number HBV clones derived from cirrhotic patients (60% of HCC-derived clones and 38% of non-HCC clones) harbored preS deletions. The preS deletions mostly affected the very 5' end of the preS1 region or the 5'/central portion of the preS2 region (Fig. 3A). Considering that the P protein moiety encoded by the preS region serves as a less important spacer (Bartenschlager and Schaller, 1988), these deletions apparently maintained P protein function. Indeed, many clones with preS deletions (3.6, 4.9, 6.1, 8.3, 8.4, 11.1, 11.2, 12.3, 13.1, 13.2, 17.1, and 19.3) were replication competent, some to very high level (Fig. 2). Deletions in the preS2 region would remove residues 131–141, 122–141, 128–143, and 137–141 in L protein (Fig. 3B, Supplementary Table 2), and corresponding sequence in M protein. Deletions in the preS2 domain affected 8 clones (2.10, 3.5, 3.6, 6.1, 13.1, 13.2, 15.3, 15.5) and remarkably, all but clone 6.1 were from HCC patients. Very similar deletions have been described from another study based on Korean patients (Ahn et al., 2010), and from pediatric HCC cases in Asia (Abe et al., 2009). In this regard, antibodies raised against preS peptides 120–145 and 133–151 could neutralize HBV infectivity in chimpanzees (Itoh et al., 1986; Neurath et al., 1986), and monoclonal antibodies against preS peptide 132–143 or recognizing peptide 132–135 prevented HBV infection in cell culture (Li et al., 2016; Ryu et al., 1997). Therefore, these deletions may escape clustered neutralizing epitopes in the preS2

domain (Fig. 3B). Based on in vitro studies using artificial constructs (Le Seyec et al., 1998; Ni et al., 2000), such deletion mutants are most likely still infectious.

Deletions at the 5' end of the preS1 region often ablated its translation initiation site but permitted initiation from the next inframe ATG codon (clones 7.9, 8.3, 8.4, 11.1, 11.2). The consequence would be loss of the N-terminal 11 residues in L protein. Such an N-terminally truncated L protein would still be myristoylated (codon 13 codes for glycine), and structurally similar to that produced by wildtype genotype D. That genotype has a 33-nt deletion at the 5' end of the preS1 region relative to other genotypes. The biological significance of deleting N-terminal 11 residues from L protein is unclear at present. However, the N-terminus of L protein represents the binding site for NTCP, the high-affinity HBV receptor (Yan et al., 2013). The preS1 sequence 21–47 has been recognized as a neutralizing epitope in chimpanzees (Neurath et al., 1989), and a neutralizing monoclonal antibody recognizes preS1 sequence 20–23 (Glebe et al., 2003) (Fig. 3B). A previous study suggested that myristoylation helps to mask the neutralizing epitopes in L protein (Bremer et al., 2011). In this connection, moving the myristoylation site downstream by 11 residues most likely makes virions inaccessible to some neutralizing antibodies generated against full-length L protein, thus representing a novel mechanism of immune escape. Nevertheless, for the 5 clones studied here, clone 7.9 had a frameshift mutation in the S domain, while clones 8.3 and 8.4 had a nonsense mutation at codon 61 in the S region. Thus, only clones 11.1 and 11.2 could express a genotype-D like L protein.

A major focus of the present study was to characterize the HBsAg expression profile of the 36 full-length HBV clones and to explore the structural bases for impaired HBsAg production or secretion. A total of 10 clones were found to be negative in HBsAg production, although the HBsAg-negative phenotype was not more prevalent in the HCC group (Supplementary Table 2). Many of these clones harbored preS deletion in addition to a nonsense mutation in the S or preS region, or possible immune escape mutations in the “a” determinant (residues 124–147) within the S domain. Site-directed mutagenesis using the 0.7mer expression construct for the three envelope proteins demonstrated that deletions affecting the 3' end rather than the 5' end or the central portion of the SPII promoter completely abolished HBsAg production (Fig. 6). Among the missense mutations analyzed (D99G, P120T, C121Y, I126S, G145R, and R160G), only D99G caused HBsAg-negative phenotype (Fig. 6). D99 is located at the second transmembrane segment of the S protein. The drastic change from aspartic acid to glycine may affect membrane translocation. The G72* and Q121* nonsense mutations in the preS region could explain the lack of L protein expression and HBsAg negative phenotype of clones 12.1, 12.3, 8.3, and 8.4, while the C69* and S61* nonsense mutations in the S region were confirmed to abolish HBsAg production (Fig. 6).

More distal nonsense mutations in the S region include W172*, W182*, W201*, and L209* (Fig. 4). The W172* mutation is associated with resistance to nucleos(t)ide analogues and secondary to substitution in the reverse transcriptase (RT) domain of the P protein (rtA181T/sW172*). That mutation has been found to abolish virion and HBsAg secretion (Warner and Locarnini, 2008), and may increase the HCC risk (Yeh et al., 2011). In our case, the sW172* mutation was silent at the P protein level and hence not selected by antiviral therapy. We

found the L209*, W182*, and especially W201* mutations markedly reduced intracellular S protein and HBsAg (Fig. 6). On the other hand, HBsAg secretion (the ratio of extracellular HBsAg/intracellular HBsAg) was completely blocked by the L209* mutation and diminished by the W182* and W201* mutations. The impact of these mutations on virion secretion warrants further investigation.

5. Conclusions

HBV genomes derived from Korean cirrhotic patients harbored a variety of mutations at the genomic level including core promoter mutations, lost preS1 or preS2 translation initiation sites, nonsense mutations in the precore region, preS/S gene, or 3' end of the core gene, large in-frame deletions in the preS region or core gene. At the protein level, many of these HBV clones were unable to express HBeAg, and some were also deficient in M protein expression, or expressed a genotype D-like L protein. Many clones failed to produce HBsAg due to deletion in the SPII promoter or premature nonsense mutation in the S or preS region. Some clones had a major antigenic epitope in the core protein removed, while switch to a genotype D-like L protein, deletions in the preS2 domain, and missense mutations in the "a" determinant of the S domain may help the virus to escape neutralizing antibodies against the preS1, preS2, and S domains, respectively. Compared with non-HCC clones, the HCC clones showed higher prevalence of T1753C and G1896A mutations, and preS deletions, especially those removing a neutralizing epitope in the preS2 domain. They were more likely to have a functional core gene and to maintain a high level of genome replication capacity. They were also more likely to have mutations affecting the C-terminal four residues of core protein, although its functional consequence remains unclear.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This study was supported by NIH grants AI103648, AI107618, AI113394, and AI116639, and also by a grant from National Science Foundation of China (81371822).

Abbreviations:

HBV	hepatitis B virus
HCC	hepatocellular carcinoma
HBsAg	hepatitis B surface antigen
P	polymerase
HBeAg	hepatitis B e antigen
L, M, S	large middle and small envelope proteins
pcRNA	precore RNA

pgRNA	pregenomic RNA
PCR	polymerase chain reaction
cp	core promoter
pc	precore
sup	supernatant
ELISA	enzyme-linked immunosorbent assay
PEG	polyethylene glycol
WT	wild type

References

- Abe K, Thung SN, Wu HC, Tran TT, Le Hoang P, Truong KD, Inoue A, Jang KK, Su IJ, 2009 Pre-S2 deletion mutants of hepatitis B virus could have an important role in hepatocarcinogenesis in Asian children. *Cancer Sci.* 100, 2249–2254. [PubMed: 19719772]
- Ahn SH, Yuen L, Han K, Littlejohn M, Chang H, Damerow H, Ayres A, Heo J, Locarnini S, Revill P, 2010 Molecular and clinical characteristics of hepatitis B virus in Korea. *J. Med. Virol.* 82, 1126–1134. [PubMed: 20513074]
- Baptista M, Kramvis A, Kew MC, 1999 High prevalence of 1762(T) 1764(A) mutations in the basic core promoter of hepatitis B virus isolated from black Africans with hepatocellular carcinoma compared with asymptomatic carriers. *Hepatology* 29, 946–953. [PubMed: 10051502]
- Bartenschlager R, Schaller H, 1988 The amino-terminal domain of the hepatitis B virus P-gene encodes the terminal protein (genome-linked protein) believed to prime reverse transcription. *EMBO J.* 7, 4185–4192. [PubMed: 2854056]
- Baumert T, Rogers S, Hasegawa K, Liang T, 1996 Two core promoter mutations identified in a hepatitis B virus strain associated with fulminant hepatitis result in enhanced viral replication. *J. Clin. Invest.* 98, 2268–2276. [PubMed: 8941643]
- Bremer CM, Sominskaya I, Skrastina D, Pumpens P, El Wahed AA, Beutling U, Frank R, Fritz H-J, Hunsmann G, Gerlich WH, Glebe D, 2011 N-terminal myristoylation-dependent masking of neutralizing epitopes in the preS1 attachment site of hepatitis B virus. *J. Hepatol.* 55, 29–37. [PubMed: 21145866]
- Carman WF, Zanetti AR, Karayiannis P, Waters J, Manziello G, Tanzi E, Zuckerman AJ, Thomas HC, 1990 Vaccine-induced escape mutant of hepatitis B virus. *Lancet* 336, 325–329. [PubMed: 1697396]
- Chen BF, Liu CJ, Jow GM, Chen PJ, Kao JH, Chen DS, 2006 High prevalence and mapping of pre-S deletion in hepatitis B virus carriers with progressive liver diseases. *Gastroenterology* 130, 1153–1168. [PubMed: 16618410]
- Chen MT, Changchien CS, Lee CM, Hung CH, Hu TH, Wang JH, Wang JC, Lu SN, 2008 Combined mutations in preS/surface and core promoter/precore regions of hepatitis B virus increase the risk of hepatocellular carcinoma: a case-control study. *J. Infect. Dis.* 198, 1634–1642. [PubMed: 18939932]
- Ehata T, Omata M, Yokosuka O, Hosoda K, Ohto M, 1992 Variations in codons 84–101 in the core nucleotide sequence correlate with hepatocellular injury in chronic hepatitis B virus infection. *J. Clin. Invest.* 89, 332–338. [PubMed: 1729279]
- Fang ZL, Sabin CA, Dong BQ, Wei SC, Chen QY, Fang KX, Yang JY, Huang J, Wang XY, Harrison TJ, 2008 Hepatitis B virus preS deletion mutation is a risk factor for hepatocellular carcinoma: a matched nested case-control study. *J. Gen. Virol.* 89, 2882–2890. [PubMed: 18931087]

- Garcia T, Li J, Sureau C, Ito K, Qin Y, Wands J, Tong S, 2009 Drastic reduction in the production of subviral particles does not impair hepatitis B virus virion secretion. *J. Virol.* 83, 11152–11165. [PubMed: 19706705]
- Glebe D, Alikbari M, Krass P, Knoop EV, Valerius KP, Gerlich WH, 2003 Pre-S1 antigen-dependent infection of *Tupaia* hepatocytes with human hepatitis B virus. *J. Virol.* 77, 9511–9521. [PubMed: 12915565]
- Ito K, Kim K-H, Lok A, Tong S, 2009 Characterization of genotype-specific carboxyl-terminal cleavage sites of hepatitis B virus e antigen precursor and identification of furin as the candidate enzyme. *J. Virol.* 83, 3507–3517. [PubMed: 19193799]
- Itoh Y, Takai E, Ohnuma H, Kitajima K, Tsuda F, Machida A, Mishihiro S, Nakamura T, Miyakawa Y, Mayumi M, 1986 A synthetic peptide vaccine involving the product of the pre-S(2) region of hepatitis B virus DNA: protective efficacy in chimpanzees. *Proc. Natl. Acad. Sci. U. S. A.* 83, 9174–9178. [PubMed: 3466181]
- Kao JH, Chen PJ, Lai MY, Chen DS, 2003 Basal core promoter mutations of hepatitis B virus increase the risk of hepatocellular carcinoma in hepatitis B carriers. *Gastroenterology* 124, 327–334. [PubMed: 12557138]
- Kim JK, Chang HY, Lee JM, Baatarkhuu O, Yoon YJ, Park JY, Kim Y, Han KH, Chon CY, Ahn S, 2009 Specific mutations in the enhancer II/core promoter/ precore regions of hepatitis B virus subgenotype C2 in Korean patients with hepatocellular carcinoma. *J. Med. Virol.* 81, 1002–1008. [PubMed: 19382267]
- Kramvis A, 2014 Genotypes and genetic variability of hepatitis B virus. *Intervirology* 57, 141–150. [PubMed: 25034481]
- Le Seyec J, Chouteau P, Cannie I, Guguen-Guillouze C, Gripon P, 1998 Role of the pre-S2 domain of the large envelope protein in hepatitis B virus assembly and infectivity. *J. Virol.* 72, 5573–5578. [PubMed: 9621015]
- Li J, Zong L, Sureau C, Barker L, Wands J, Tong S, 2016 Unusual features of sodium taurocholate cotransporting polypeptide as a hepatitis B virus receptor. *J. Virol.* 90, 8302–8313. [PubMed: 27384660]
- Milich DR, McLachlan A, Moriarty A, Thornton GB, 1987 Immune response to hepatitis B virus core antigen (HBcAg): localization of T cell recognition sites within HBcAg/HBeAg. *J. Immunol.* 138, 1223–1231.
- Nassal M, 1992 The arginine-rich domain of the hepatitis B virus core protein is required for pregenome encapsidation and productive viral positive-strand DNA synthesis but not for virus assembly. *J. Virol.* 66, 4107–4116. [PubMed: 1602535]
- Neurath A, Kent S, Parker K, Prince A, Strick N, Brotman B, Sproul P, 1986 Antibodies to a synthetic peptide from the preS 120–145 region of the hepatitis B virus envelope are virus neutralizing. *Vaccine* 4, 35–37. [PubMed: 2421497]
- Neurath A, Seto B, Strick N, 1989 Antibodies to synthetic peptides from the preS1 region of the hepatitis B virus (HBV) envelope (env) protein are virus -neutralizing and protective. *Vaccine* 7, 234–236. [PubMed: 2476893]
- Ni Y, Sonnabend J, Seitz S, Urban S, 2000 The pre-S2 domain of the hepatitis B virus is dispensable for infectivity but serves a spacer function for L-protein-connected virus assembly. *J. Virol.* 74, 3879–3888.
- Norder H, Courouce AM, Coursaget P, Echevarria JM, Lee SD, Mushahwar IK, Robertson BH, Locarnini S, Magnius L, 2004 Genetic diversity of hepatitis B virus strains derived worldwide: genotypes, subgenotypes, and HBsAg subtypes. *Intervirology* 47, 289–309. [PubMed: 15564741]
- Orito E, Mizokami M, Sakugawa H, Michitaka K, Ishikawa K, Ichida T, Okanoue T, Yotsuyanagi H, Iino S, 2001 A case-control study for clinical and molecular biological differences between hepatitis B viruses of genotypes B and C. Japan HBV Genotype Research Group. *Hepatology* 33, 218–223. [PubMed: 11124839]
- Parekh S, Zoulim F, Ahn SH, Tsai A, Li J, Kawai S, Khan N, Trepo C, Wands JR, Tong S, 2003 Genome replication, virion secretion, and e antigen expression of naturally occurring hepatitis B virus core promoter mutants. *J. Virol.* 77, 6601–6612. [PubMed: 12767980]

- Ryu C, Gripon P, Park H, Park S, Kim Y, Guguen-Guillouzo C, Yoo O, Hong H, 1997 *In vitro* neutralization of hepatitis B virus by monoclonal antibodies against the viral surface antigen. *J. Med. Virol.* 52, 226–233. [PubMed: 9179773]
- Tong S, Revill P, 2016 Overview of hepatitis B viral replication and genetic variability. *J. Hepatol.* 64 (Suppl. 1), S4–S16. [PubMed: 27084035]
- Tong S, Li J, Wands J, Wen Y, 2013 Hepatitis B virus genetic variants: biological properties and clinical implications. *Emerg. Microb. Infect.* 2, e10.
- Tsai A, Kawai S, Kwei K, Gewaily D, Hutter A, Tong D, Li J, Wands J, Tong S, 2009 Chimeric constructs between two hepatitis B virus genomes confirm transcriptional impact of core promoter mutations and reveal multiple effects of core gene mutations. *Virology* 387, 364–372. [PubMed: 19327810]
- Warner N, Locarnini S, 2008 The antiviral drug selected hepatitis B virus rtA181T/ sW172* mutant has a dominant negative secretion defect and alters the typical profile of viral rebound. *Hepatology* 48, 88–98. [PubMed: 18537180]
- Yan H, Zhong G, Xu G, He W, Jing Z, Gao Z, Huang Y, Qi Y, Peng B, Wang H, Fu L, Song M, Chen P, Gao W, Ren B, Sun Y, Cai T, Feng X, Sui J, Li W, 2013 Sodium taurocholate cotransporting polypeptide is a functional receptor for human hepatitis B and D virus. *Elife* 1, e00049.
- Yeh CT, Chen T, Hsu CW, Chen YC, Lai WM, Liang KH, Chen TC, 2011 Emergence of the rtA181T/ sW172* mutant increased the risk of hepatoma occurrence in patients with lamivudine-resistant chronic hepatitis B. *BMC Cancer* 11, 398. [PubMed: 21933446]
- Yuen MF, Sablon E, Yuan HJ, Wong DK, Hui CK, Wong BC, Chan AO, Lai CL, 2003 Significance of hepatitis B genotype in acute exacerbation, HBeAg seroconversion, cirrhosis-related complications, and hepatocellular carcinoma. *Hepatology* 37, 562–567. [PubMed: 12601354]

Fig. 1.

Comparison of mutation rates and functional properties between HCC and non-HCC-derived clones. The mutations analyzed include core promoter mutations A1762T/G1764A and T1753C, G1896A precore mutation, core gene defect (lost initiation site, large internal deletion, or premature nonsense mutation), internal deletions in the preS region, nonsense (NS) or frameshift (FS) mutation in the preS region, mutated preS2 ATG to cause M-minus phenotype, as well as nonsense (NS) or frameshift (FS) mutation in the S gene. Functional properties include percentage of clones with high replication capacity, lack of HBeAg expression, or lack of HBsAg expression. Please note that the mutation rates for G1896A and core gene defect are based on partial sequencing of 79 clones (see Table 2 for details),

whereas the rates for other mutations and functional properties are based on 36 clones (part of the 79 clones) available as SphI dimer.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Fig. 2.

Replication capacities of full-length HBV clones and impact of back mutations in precore and/or core promoter. (A& B) Replication capacities of 36 clones were compared in Huh7 cells (A) and HepG2 cells (B). Cells were harvested 5 days post-transfection. Cloned HBV DNA cut with restriction enzymes served as size marker, while the mock transfected cells served as a negative control. (C) Role of core promoter and precore mutations in the high replication capacity of clone 4.2 and clone 19.3. CP + PC: original clone with both types of mutations. PC: a construct retaining just the precore mutation. CP: a construct retaining just the core promoter mutations. WT: a construct with both precore and core promoter sequences reverted to wild type. Cloned HBV DNA (100 pg and 20 pg) served as size marker.

Fig. 3.

Large in-frame deletions in the preS region and their impact on protein translation. The data are based on complete sequencing of the 36 HBV clones available as SphI dimer. (A) Deletions at the nucleotide level. Shown on top is the entire preS1 region (2848–3204) and part of the preS2 region, with the initiating ATGs shown as triangles. The SPII promoter is shown as an open box. Each deletion is indicated by a line, followed by numbers indicating the first and last nucleotides of the deleted sequence. The HBV clone(s) with such a deletion is provided inside parenthesis. (B) Deletions at the amino acid level. Shown on top is the entire preS1 domain (residues 1–119) and part of the preS2 region (starting at residue 120). The methionine residue at position 12 is shown as an open triangle. In case of deletions removing the first ATG in the preS1 region, a shortened L protein without the first 11 residues would be expressed. Open boxes indicate neutralizing epitopes in the preS1 domain and preS2 domain, respectively.

Fig. 4.

Mutations in the preS/S gene capable of truncating L/and or S protein or preventing M protein expression. Shown on top are the coding sequences for L, M, and S proteins (with 400, 281, and 226 residues), respectively. A solid triangle indicates translation initiation site. For the mutations found in this study, a star denotes a nonsense mutation (a stop codon). Please note that nonsense mutations in the S region are numbered according to S region itself, rather than counting from the preS1 region. FS188 means a frameshift mutation at codon 188 in the S region due to a single-nucleotide insertion. The M120X mutation denotes any single nucleotide change at the preS2 ATG codon, which would prevent the translation of M protein.

Fig. 5.

Comparison of envelope protein expression and HBsAg production among 36 full-length HBV clones. Huh7 cells were transiently transfected with SphI dimer of the 36 HBV clones and harvested 4 days later. (A) Western blot analysis of intracellular L protein and S protein, using beta-actin as a loading control. (B) ELISA for HBsAg in cell lysate and culture supernatant. Cells were lysed in 80 μ l of lysis buffer, and 20 μ l was used for Western blot. For HBsAg measurement, cell lysate was diluted 1:600 while culture supernatant was diluted 1:500. The Y-axis denotes OD450 value of HBsAg.

Fig. 6.

Impact of deletions in the preS region, amino acid substitutions in the S domain, and nonsense mutations in the S gene on S protein expression and HBsAg production. Mutations were introduced to a 0.7mer expression construct for HBsAg, with the parental construct (WT) transfected in parallel to serve as a positive control. (A) Western blot analysis of intracellular S protein, using beta-actin as a loading control. (B) ELISA for HBsAg in cell lysate and culture supernatant after 1:600 and 1:500 dilution, respectively. The preS

deletions (d) shown are at the nucleotide level. FS189: a frameshift mutation due to a single nt insertion at position 718. The Y-axis denotes OD450 value of HBsAg.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

Features of Korean HCC patients and their matched controls.

patient ^a	sex	age	Liver cirrhosis	HCC	HBV DNA (pg/ml)	HBeAg	HBsAg
1	M	43	+	+	35.2	0.096	1.720
6	M	43	+	–	1461	1.656	1.804
2	M	48	+	+	823	0.296	1.747
7	M	43	+	–	523	0.495	1.800
3	F	58	+	+	6000	0.485	1.758
8	F	56	+	–	453	0.149	1.790
4	M	57	+	+	61.9	0.506	1.791
9 ^b	M	55	+	–	130	N.D.	N.D.
5	M	46	+	+	19.8	0.092	1.785
10 ^c	M	45	+	–	2120	N.D.	N.D.
11	M	65	+	+	522.3	0.130	1.752
12	M	63	+	–	2525	2.266	0.651
13	M	53	+	+	14.9	0.100	1.908
14	M	48	+	–	736	2.307	1.904
15	M	57	+	+	65.6	1.238	1.885
16	M	54	+	–	203	2.346	1.882
17	M	52	+	+	3.4	0.126	1.857
18	M	50	+	–	1.8	0.106	1.840
19	M	36	+	+	21	0.098	1.649
20	M	39	+	–	139	0.325	1.751

^aHCC patients shown in bold.^b#9 sample failed to be amplified by PCR.^cNone of the 5 clones from #10 has a functional core gene. M: male; F: female; N.D.: not determined.

Table 2

Precore and core gene mutations found in 41 non-HCC clones and 38 HCC clones.

	Non-HCC clones	HCC clones
G1896A in precore region	8 clones from 4 patients (7.7, 16.2, 18.1, 18.2, 18.3, 18.4, 20.3, 20.5)	21 clones from 8 patients (1.2, 1.3, 1.4, 1.5, 1.6, 2.6, 2.10, 4.2, 4.5, 4.6, 4.7, 4.9, 5.1, 11.1, 11.2, 13.1, 13.2, 17.1, 17.2, 17.6, 19.3)
mutated initiation codon in core gene	1 clone (7.10:T1902C)	0
frameshift (FS) in core gene	1 clone (7.3:FS after codon 20)	1 clone (4.4:FS after codon 18)
large internal deletion in core gene	6 clones from 3 patients (10.1 and 10.2: codon 89–128 removed; 12.2: codon 88–128 removed; 20.1, 20.4, and 20.6: codon 79–107 removed)	1 clone (5.2:codon 84–121 removed)
Large C-terminal deletion of core protein	6 clones from 2 patients (8.1, 8.2, and 8.5: truncation after aa35; 10.3:truncation after aa41; 10.4 and 10.5:truncation after aa 87)	1 clone (4.7)
E180G in core protein	0	4 clones from 1 patient (4.2, 4.5, 4.6, 4.9)
S181P in core protein	2 clones from 1 patient (16.1, 16.3)	7 clones from 3 patients (3.2, 3.3, 3.5, 3.6, 13.1, 13.2, 15.5)
Q182K in core protein	1 clone (14.3)	1 clone (5.1)
Q182* (deletion of C-terminal 2aa) in core protein	1 clone (16.2)	8 clones from 2 patients (2.1, 2.6, 2.7, 2.8, 2.9, 17.1, 17.2, 17.6)

Table 3

Clones with markedly altered HBx structure or lost expression.

clone ^a	mutation
2.10	154-aa HBx extended by 24 aa due to deletion of nt 1834.
3.5	8-aa insertion between residues 93 and 94 due to 24-nt insertion.
4.9	mutated translation initiation site.
11.2	N-terminal 150 aa of HBx fused to 31aa due to deletion of nt 1821. Extra sequence similar to that of Kca2.10.
12.1	N-terminal 118aa of HBx fused to 10 aa due to deletion of nt 1727.
12.3	N-terminal 118aa of HBx fused to 10 aa due to deletion of nt 1727.

^aShown in bold are HCC clones.

Table 4

Mutations responsible for the lack of L and/or S protein expression.

Clone ^a	Lack of L protein	Lack of S protein
4.9	deletion of preS aa 42–78 + truncation after aa 181 in S domain	deletion of nt 2973–3083 + truncation after aa 181
5.2	truncation after aa 68 in S domain	truncation after aa 68
7.9	fusion with P protein after aa 188 in S domain	fusion with P protein after aa 188
8.3	truncation after aa 120 in preS domain	truncation after aa 60
8.4	truncation after aa 120 in preS domain	truncation after aa 60
12.1	truncation after aa 71 in preS domain	I126S + truncation after aa 200
12.3	truncation after aa 71 in preS domain	NA ^b
14.3	NA ^b	D99G
15.5	deletion of preS aa 43–79 and 137–141, truncation after aa 181 in S domain	deletion of nt 2973–3083 + truncation after aa 181
17.1	NA ^b	deletion of nt 3123–3164
19.3	1st preS1 ATG lost by deletion, 2nd ATG lost by point mutation	deletion of nt 3132–60

^aHCC clones are shown in bold.

^bNA: not applicable.