**Name of journal: World Journal of Gastroenterology**

**Manuscript No: 38667**

**Manuscript Type: REVIEW**

**Hepatitis B virus pre-S/S variants in liver diseases**

Chen BF. HBV pre-S/S mutants

Bing-Fang Chen

Bing-Fang Chen, School of Medicine, Fu-Jen Catholic University, New Taipei City 24205, Taiwan

**ORCID number: (**0000-0002-3697-1533).

**Author contributions:** Chen BF wrote the paper.

**Supported by** the grant from the National Science Council (NSC 96-2320-B-030-004-MY3), Executive Yuan, Taiwan.

**Conflict-of-interest statement:** No potential conflicts of interest.

**Open-Access:** This article is an open-access article which was selected by an in-house editor and fully peer-reviewed by external reviewers. It is distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

**Manuscript source:** Invited manuscript

**Correspondence to: Bing-Fang Chen, PhD, Professor,** School of Medicine, Fu-Jen Catholic University, 510 Chung-Cheng Road, Hsin-Chuang Dist, New Taipei City 24205, Taiwan. nurs1018@mail.fju.edu.tw

**Telephone:** +886-2-29053428

**Fax:** +886-2-29052096

**Received:** March 8, 2018

**Peer-review started:** March 9, 2018

**First decision:** March 21, 2018

**Revised:** March 29, 2018

**Accepted:** March 31, 2018

**Article in press:**

**Published online:**

**Abstract**

Chronic hepatitis B is a global health problem. The clinical outcomes of chronic hepatitis B infection include asymptomatic carrier state, chronic hepatitis (CH), liver cirrhosis (LC), and hepatocellular carcinoma (HCC). Because of the spontaneous error rate inherent to viral reverse transcriptase, the hepatitis B virus (HBV) genome evolves during the course of infection under the antiviral pressure of host immunity. The clinical significance of pre-S/S variants has become increasingly recognized in patients with chronic HBV infection. Pre-S/S variants are often identified in hepatitis B carriers with CH, LC, and HCC, which suggests that these naturally occurring pre-S/S variants may contribute to the development of progressive liver damage and hepatocarcinogenesis. This paper reviews the function of the pre-S/S region along with recent findings related to the role of pre-S/S variants in liver diseases. According to the mutation type, five pre-S/S variants have been identified: pre-S deletion, pre-S point mutation, pre-S1 splice variant, C-terminus S point mutation, and pre-S/S nonsense mutation. Their associations with HBV genotype and the possible pathogenesis of pre-S/S variants are discussed. Different pre-S/S variants cause liver diseases through different mechanisms. Most cause the intracellular retention of HBV envelope proteins and induction of endoplasmic reticulum stress, which results in liver diseases. Pre-S/S variants should be routinely determined in HBV carriers to help identify individuals who may be at a high risk of less favorable liver disease progression. Additional investigations are required to explore the molecular mechanisms of the pre-S/S variants involved in the pathogenesis of each stage of liver disease.

**Key words:** hepatitis B virus; pre-S/S mutant; pre-S deletion; splice variant; spPS1; chronic hepatitis; liver cirrhosis; hepatocellular carcinoma

**© The Author(s) 2018.** Published by Baishideng Publishing Group Inc. All rights reserved.

**Core tip:** Naturally occurring hepatitis B virus (HBV) pre-S/S variants have been identified and associated with progressive liver diseases. In this review, the author discusses five pre-S/S variants: pre-S deletion, pre-S point mutation, pre-S1 splice variant, C-terminus S point mutation, and pre-S/S nonsense mutation. Their associations with HBV genotype and the possible pathogenesis of pre-S/S variants are also discussed. Different pre-S/S variants cause liver diseases through different mechanisms. Most cause the intracellular retention of HBV envelope proteins and induction of endoplasmic reticulum stress, resulting in liver diseases. The exact pathogenesis of pre-S/S variants requires further investigation.

Chen BF. Hepatitis B virus pre-S/S variants in liver diseases. *World J Gastroenterol* 2018; In press

**INTRODUCTION**

Hepatitis B virus (HBV) infection, which causes acute and chronic liver diseases, is a global health concern. The majority of acute HBV infections are self-limited, whereas chronic HBV infection usually results in a lifelong course. Chronic HBV infection can result in numerous clinical conditions, including asymptomatic HBV carrier (ASC), chronic hepatitis (CH), liver cirrhosis (LC), and hepatocellular carcinoma (HCC)[1,2]. More than 350 million people worldwide are estimated to have chronic HBV infection, and more than 25% of the chronically infected patients in Asia die because of HBV-related chronic diseases. The outcomes of HBV infection vary, which is likely because of differences in the host and viral factors.

To date, 10 HBV genotypes, designated as genotypes A to J, have been identified based on a divergence of >8% over the entire genomic sequence. These 10 HBV genotypes are distributed in specific geographical locations[3,4]. Genotypes A (HBV/A) and D (HBV/D) are prevalent in Africa, Europe, and the Americas; genotypes B (HBV/B) and C (HBV/C) in Asia; genotype E (HBV/E) in sub-Saharan Africa; genotypes F and H in Southern and Central America; genotype G in France, Germany, and the United States; genotype I in Vietnam and Laos; and genotype J in Japan’s Ryukyu islands. All genotypes can lead to progressive liver disease, but the clinical implications of each genotype differ. For example, patients infected by the HBV/C or HBV/D strain have a higher frequency of basal core promoter mutations, a lower response rate to interferon therapy, and a more rapid progression to liver fibrosis and HCC than those infected by the HBV/B or HBV/A strain[3,4]. In addition, carriers infected by HBV/C have a higher rate of pre-S deletions than those infected by HBV/B[5,6]. Collectively, these data indicate pathogenic and therapeutic differences among the HBV genotypes[3,4].

HBV is a small (42 nm) enveloped DNA virus, whose genome consists of partially double–stranded circular DNA that is 3182-3248 bp in length (varying with the genotype). Four genes — pre-S/S, precore (PC)/core (C), Pol, and X — encode seven polypeptides, including the structural proteins of the virion envelope and core, a small transcriptional transactivator, and a large polymerase protein with reverse transcriptase (RT) and RNase H (RH) activity (Figure 1). The pre-S/S gene has three in-frame initiation codons and encodes the small (S) envelope proteins as well as the middle (M) and large (L) envelope proteins, which contain pre-S2 and pre-S (pre-S1 and pre-S2) sequences, respectively (Figure 2A). The PC/C gene has two in-frame initiation codons and encodes the core antigen plus HBe protein, which is processed to produce soluble hepatitis B e antigen[1]. HBV replicates through the reverse transcription of an RNA intermediate, but because the RT lacks a proofreading function, errors in HBV DNA replication occur at a much higher rate than for other DNA viruses. The estimated rate of nucleotide substitution is approximately 1.4-3.2 × 10-5 per site per year**[**7]. These naturally occurring mutants evolve during the course of infection under the antiviral pressure of the host immune system or exogenous factors, including immunization or specific therapy[8]. Such HBV mutants display alteration of epitopes vital to host immune recognition, enhanced virulence with increased replication of HBV, and resistance to antiviral therapies while facilitating cell attachment or penetration[9,10]. These viral mutants, including basal core promoter, PC mutation, pre-S deletion, pre-S mutation, S mutants, and splice variants[5,6,11-27], have been associated with an increased risk of liver diseases.

The clinical significance of these naturally occurring mutants has become increasingly recognized in patients with both acute and chronic HBV infections[8-10,21,26,27]. In this article, the function of the pre-S/S region and recent findings related to the role of pre-S/S variants on liver diseases is discussed and reviewed.

**THE BIOLOGICAL FUNCTION OF THE PRE-S/S REGION**

The pre-S/S gene has three open reading frames (ORFs) that encode three forms of hepatitis B surface antigen (HBsAg): the L, M, and S structural proteins of the viral envelope. However, these proteins are translated from different mRNAs: the L protein is translated from a long 2.4 kb pre-S1 RNA transcript, whereas the M and S proteins are translated from a slightly shorter 2.1 kb S RNA transcript (Figure 2A). The S protein consists of 226 amino acids (aa). The M protein is an extension of the S protein, with an additional 55 aa (*i.e.*, pre-S2 region). The L protein is an extension of the M protein, with an additional 108–119 aa depending on the genotype (i.e., pre-S1 region). The aa sequence present at the C termini of the L and M proteins is identical to the S protein and is referred to as the S region. The pre-S (pre-S1 and pre-S2) region of the L protein is crucial for viral replication. It contains several functional sites: the hepatocyte binding site, which is essential for the attachment of HBV to liver cells; the S promoter and the CCAAT binding factor binding site, which is essential for S RNA transcription; the heat-shock protein 70 (Hsc70) binding site and the cytosolic anchorage determinant (CAD), which are essential for the dual topology (T) of L proteins; the nucleocapsid binding site (NBS), which is essential for virion morphogenesis; the site for viral secretion (VS); and the site for polymerized human albumin (pHSA) (Figure 3)[28-33].The pre-S region also plays an essential role in the interaction with the immune responses because it contains both B- and T-cell epitopes (Figure 3)[34-39]. By contrast, the biological role of M protein in the viral life cycle has been controversial. *In vitro* studies have suggested that M protein is not essential for viral replication, virion morphogenesis, or infectivity. Huang *et al* defined a novel regulatory role for M protein, which may undergo a proteolytic process to generate an MHBsau (aa 1-57 of M protein) species to upregulate the transcription of S promoter[40]. In addition, the pre-S2 region of M protein binds to pHSA (aa 3-16), but the significance of this binding is unknown[34]. The S proteins are required for virion morphogenesis and secretion, and they also contain both B- and T-cell epitopes[26,41].

HBV envelope proteins are synthesized at the endoplasmic reticulum (ER). HBV envelope proteins have an unusual feature; they have multiple transmembrane domains that span the ER with loops of amino acids internal and external to the cytosol (Figure 4A)[41]. The S protein spans the ER membrane through four transmembrane domains(TM 1-4) that are linked by internal and external loops[41]. The loop of amino acids linking TM2 and TM3 is external to the ER and comprises aa 99–169. This loop is known as the "a”determinant (aa 122-148), and it is of vital virological and clinical signiﬁcance as it is a major antigenic determinant of HBV. The transmembrane topology of the M protein is identical to the S protein. By contrast, the L protein has two transmembrane topologies. On biogenesis, the CAD of the pre-S1 region interacts with the cognate heat-shock protein Hsc70, thereby preventing cotranslational pre-S translocation to remain the pre-S domain of L cytosolic[42,43]. During maturation, approximately half of the L molecules posttranslationally translocate their pre-S region into the ER, thereby generating a dual topology (Figure 4A)[41-43]. The L protein serves its topological opposing functions in the virus life cycle by orientating the pre-S domain at both the cytosolic (i-Pre-S, inside the virus) and luminal (e-Pre-S, outside the virus) locations, i-Pre-S for capsid envelopment and e-Pre-S for receptor binding[41].

**ASSOCIATION BETWEEN HBV PRE-S/S VARIANTS AND LIVER DISEASES**

Owing to the spontaneous error rate of viral reverse transcription, naturally occurring HBV mutants arise during the course of a patient’s infection under the pressure of host immunity or speciﬁc therapy[8]. Recently, many investigations have reported that pre-S/S variants are associated with the development of liver diseases[5,6,11-14,26-27]. Here, according to the mutation type, five pre-S/S variants—pre-S deletion, pre-S point mutation, pre-S1 splice variant, C-terminus S point mutation, and pre-S/S nonsense mutation—are reviewed. The pre-S region is the most variable sequence of the viral genome and changes with the genotype. The HBV genotype may influence the emergence of different pre-S variants; thus, it is also reviewed.

***Pre-S deletion and genotype***

Many studies have demonstrated that pre-S deletions are associated with progressive liver diseases[5,6,11-14,26-27]. Pre-S deletion is frequently found at the C-terminal half (aa 58-119) of pre-S1 and the N-terminus (aa 1–23) of pre-S2. Most are in-frame deletions[6,11-14,26-27]. Mapping of the pre-S region has revealed that all deletion regions encompassed T- and B-cell epitopes, and most of them lost one or more functional sites, including the S promoter, T site, NBS, start codon of M, VS site, and pHSA site[6,44]. Most reports have focused on the relationship between pre-S deletion and HCC and have indicated that pre-S2 deletion is associated with HCC development in adults[5,6,11-14,21,26,27,44]. Two reports have revealed a high prevalence of HBV pre-S deletion mutation, with the mutation being recognized in 27 of 30 (90%) and 9 of 19 (47.4%) examined childhood cases of HCC[45,46]. Pre-S2 deletion also occurred frequently (20/27, 74%; 8/9, 88.8%)[45,46]. Other studies have reported a high rate of pre-S1 deletion in HBV/C-infected HCC cases[47,48]. These differences might result from the prevalence of different genotypes (or subgenotypes) in different countries. Biswas *et al* investigated the association of types of pre-S mutations with HBV genotypes from 25 cases and revealed that pre-S1 deletion (5/9, 55.56%) was common in HBV/D, pre-S2 start codon mutation (5/9, 55.56%) was frequent in HBV/A, and pre-S2 deletions (3/7, 42.85%) were frequent in HBV/C[49]. Recently, we enrolled 43 HBV/B and 43 HBV/C-infected carriers with pre-S deletion to examine the prevalence of different pre-S deletions and their associations with HBV genotypes[50]. The results showed the frequencies of some types of pre-S deletion differed between the HBV/B and HBV/C groups, whereas the frequencies of other types of pre-S deletion were similar in both genotypes[50]. Sequence alignment analysis indicated that both genotypes possessed a high frequency of deletion in the C-terminus half of the pre-S1 region and N-terminus of the pre-S2 region (86.0% and 79.1% in the HBV/B group; 69.8% and 72.1% in the HBV/C group, respectively). Epitope mapping revealed that deletion in several epitope sites was frequent in both genotypes, particularly pS1-BT and pS2-B2. Conversely, the frequency of pS2-B1 deletion was significantly higher in the HBV/B group (72.1% *vs* 37.2%, *P* = 0.002), and the frequency of pS2-T deletion was significantly higher in the HBV/C group (48.8% *vs* 25.6%, *P* = 0.044). Functional mapping revealed that the frequency of deletion in three functional sites (NBS, the start codon of M, and VS site) located in the border between the pre-S1 and pre-S2 region (aa 103–127) was significantly higher in the HBV/B group (*P* < 0.05). One variety of N-terminus pre-S1 deletion mutation demonstrating deletion of the start codon of the L protein was frequently observed in the HBV/C group (20.9% *vs* 9.3%, *P* = 0.228). The correlation of different pre-S deletion with the HBV genotype was further examined according to different clinical outcomes. Significant differences were observed between the HBV/B- and HBV/C-infected patients with LC-HCC. Deletion in the N-terminus of the pre-S2 region — including two epitope sites (pS2-B1 and pS2-B2) and three functional sites (the start codon of M, VS, and pHSA) — was significantly more frequent in the HBV/B-infected LC-HCC patients (*P* < 0.05). In Asia, HBV/B and HBV/C commonly coexist. However, their distribution differs by country[3,4]. Pre-S2 deletion has been associated with the development of HCC in Taiwan[27,45,51]. This finding may be due to HBV/B being more prevalent than HBV/C in Taiwan. HBV/C is predominant in Korea, where the N-terminus pre-S1 deletion mutant with deletion of the start codon of the L protein has been correlated with the development of HCC[47,48]. These results indicate that the tendency of different pre-S deletion varies across HBV genotypes. Therefore, the difference in genotype (or subgenotype) prevalence in different countries may influence the pattern of pre-S deletion associated with HCC.

The association of specific types of pre-S deletion with CH and LC development remains unknown. Our studies have revealed that deletion in the C-terminal half of the pre-S1 region is frequent among CH and LC patients[25,50], which is in contrast to HCC patients, who demonstrated a signiﬁcantly higher frequency of deletion in the pre-S2 region[11-14,26,27,44-46]. Functional mapping showed that deletion in the S promoter was signiﬁcantly frequent in CH and LC patients compared with that in ASCs[25,50]. The correlation among different pre-S deletion mutants with HBV genotypes in CH and LC patients was investigated, and deletion in the S promoter and the C-terminal half of pre-S1 was frequently observed in both genotypes[50]. In the CH patients, deletion in the pS1-BT and four functional sites (S promoter, Hsc70, CAD, and NBS), which are located in the C-terminal half of the pre-S1 region, was frequent in both genotypes. Conversely, deletion in the pHSA was more frequent in the HBV/B group than in the HBV/C group (88.9% *vs* 36.4%, *P* = 0.028). In the LC patients, no significant differences were observed between the HBV/B and HBV/C groups, except that deletion in the start codon of L was more frequent in the HBV/C group (42.9% *vs* 12.5%, *P* = 0.193)[50].

To understand the characteristics of these pre-S deletion mutants, five naturally occurring pre-S deletion mutants — namely one pre-S1 C-terminus half deletion mutant (dps1), two pre-S1/2 deletion mutants with deletion spanning pre-S1 and pre-S2 (dpS12a and dpS12b), and two pre-S2 deletion mutants (dpS2a and dpS2b) — were analyzed *in vitro*[52]. Functional analyses indicated that they could be divided into two groups: S promoter (dpS1 and dpS12a) and non-S promoter (dpS12b, dpS2a, and dpS2b) deletion mutants. Northern blot analysis revealed that S RNA could be transcribed in non–S-promoter deletion mutants and that the ratio of pre-S1 RNA to S RNA was similar to that in wild-type (WT) HBV transfected cells (Figure 2). Conversely, S promoter deletion mutants could not transcribe S RNA efficiently and had a higher level of pre-S1 RNA, causing an inverse ratio of pre-S1 RNA to S RNA (Figure 2). Western blot and ELISA analyses indicated that intracellular accumulation of envelope proteins was present in all pre-S deletion mutant transfected cells, especially in the S promoter deletion mutant transfected cells. Immunofluorescence analysis revealed that the mutant L proteins, unlike the WT L proteins, exhibited granular staining in the S promoter deletion variants and a perinuclear staining pattern in the non–S-promoter deletion variants[52]; other studies have reported similar findings[12,26,27,53-56]. Two types of ground glass hepatocytes (GGHs) have been deﬁned and associated with liver diseases in chronic HBV infection[27]. These GGHs contain pre-S deletion mutants that are accumulated in the ER and induce ER stress. Type I GGHs that harbor pre-S1 deletion variants display a globular or inclusion-like immunostaining pattern of HBsAg and are typical of the high viral-replicative phase of chronic HBV infection. Type II GGHs that harbor pre-S2 deletion variants with or without point mutations at the start codon of M proteins demonstrate marginal staining patterns of HBsAg, are distributed in large clusters because of their higher proliferative activity, and are characteristic of the advanced stages of chronic liver diseases[27]. Pre-S deletion mutants have been observed to induce the ER stress response, leading to the enhanced expression of vascular endothelial growth factor-A and the activation of Akt/mammalian target of rapamycin signaling in GGHs[57]. In addition,pre-S2 deletion mutants may elicit the aberrant cyclin A expression and centrosome overduplication through ER stress induction and result in cell cycle progression, cell proliferation, and anchorage-independent growth[58-60]. In addition to the induction of ER stress signals, pre-S2 deletion L proteins may directly interact with the Jun activation domain-binding protein 1, thus triggering cyclin-dependent kinase inhibitor p27 degradation, retinoblastoma hyperphosphorylation, and cell cycle progression[61]. These studies all suggest that pre-S deletion mutants may cause the intracellular retention of HBV envelope proteins, resulting in liver diseases.

***Pre-S point mutant and genotype***

The pathogenic role of pre-S point mutation has been the subject of fewer studies. Chen *et al* reported that, compared with control patients, patients with HCC had higher frequencies of pre-S deletions and amino acid substitutions at codon 4 (W4P/R), 7 (K7T/N), and 81 (A81T) in the pre-S1 regions; and at the start codon (M1V/I/A) in the pre-S2 regions[62]. By contrast, patients had a lower frequency of amino acid substitution at codon2 (Q2K/R) in the pre-S2 regions compared with control patients. The correlation between different pre-S point mutation with HBV genotype was further examined; compared with patients with HBV/B infection, patients with HBV/C infection were found to have higher frequencies of amino acid substitutions at codon 4 (17 of 79 *vs* 0 of 159; *P* < 0.001), codon 7 (14 of 79 *vs* 3 of 159; *P* < 0.001), and codon 81 (16 of 79 *vs* 2 of 159; *P* < 0.001) inpre-S1 genes[62]. Zhang *et al* also reported that compared with the HCC-free group, higher frequencies of pre-S deletions and point mutations at 11 codons — 4, 27, 51, 54, 60, 62, 100, 125, 137, 166, and 167 — were observed in the HCC group (*P* < 0.05) with either HBV/B or HBV/C[44]. Multiple logistic regression analysis revealed that pre-S deletions and point mutations at codon 51 and 167 were independent factors associated with HCC. Longitudinal observation indicated that pre-S deletions and the majority of the 11 HCC-associated pre-S point mutations existed at least 10 years before HCC development, and they were more prevalent preceding HCC development in patients from the HCC groups than the HCC-free group[44]. Five amino acid sites (codon 27, 35, 54, 137, and 167) that were under positive selection pressure were identified in the HBV/C sequences, whereas no positive selection codon was detected for HBV/B[44]. Zhang *et al*[63]later used deep sequencing to examine the dynamics of HBV quasi-species and their relationship to HCC development. In total, 32 chronic hepatitis B (CHB) patients with HCC (HCC group) and 32 matched controls were recruited[63]. HCC patients were found to have a higher intrapatient prevalence of pre-S deletions and point mutations at codons 4, 27, and 167 compared with the control patients (all *P* < 0.05). Longitudinal observation in the sera of 14 HCC patients determined that quasi-species complexity (*P* = 0.027 and 0.024 at the nucleotide level and the amin oacid level, respectively) and diversity (*P* = 0.035 and 0.031 at the nucleotide level and the amino acid level, respectively) increased as the disease progressed to HCC[63]. Another study in patients with either HBV/B or HBV/C indicated that point mutation C2964A, A2962G, and C3116T in the pre-S1 region; C7A and T53C in the pre-S2 region; and pre-S2 start codon mutation are associated with an increased risk of HCC, and a novel mutation C105T in the pre-S2 region is inversely associated with the risk of HCC[64]. Functional studies investigating pre-S point mutants have been conducted. Mun *et al*[65] demonstrated that amino acid substitution F141L in the pre-S2 region increases the risk of HCC in HBV/C-infected subjects. An *in vitro* study demonstrated that F141L-LHBs can induce cell cycle progression by down-regulating the p53 and p21 pathways and up-regulating cyclin-dependent kinase 4 and cyclin A. In a colony-forming assay, the colony-forming frequencies in cell lines expressing F141L-LHBs were approximately twice as high as those of the WTs[65]. This suggests that F141L-LHBs may have a vital role in the pathogenesis of HCC by inducing cell proliferation and transformation[65]. Zhang *et al*[63]proposed that these pre-S point mutants may cause imbalanced envelope protein production and intracellular retention of HBsAg, leading to ER stress and tumorigenesis. These studies were conducted in patients infected with HBV/B or HBV/C. Additional studies are required to evaluate whether these mutations exist in other HBV genotypes and whether the conclusions of previous studies are valid.

***Pre-S1 splice variant and genotype***

RNA splice donor and acceptor sites can be detected throughout the HBV genome. Thus, RNA splicing can occur and involve deletions of nucleotides at speciﬁc sites. To date, 14 types of spliced HBV genomes have been identiﬁed and isolated from the sera and liver tissues of HBV-infected patients[23,24,66,67]. Different introns are removed in different splicing variants, and the splicing variants vary by genotype. The splice sites of the HBV genome are not random: the five common splice donor sites are at nucleotide positions 2067, 2447, 2471, 2985, and 2087, and the five common splice acceptor sites are at nucleotide positions 489, 2350, 2236, 2902, and 282 (these nucleotide positions are based on HBV/D). These variants can be reverse transcribed and packaged with the help of WT virus to provide the necessary proteins[68,69]. Several studies have reported that spliced HBV variants enhance WT virus replication in patients with CH; these variants have been associated with advanced liver disease[23,24]. The most frequently detected splice variant, SP1, can encode a novel protein — the hepatitis B spliced protein — which has been associated with viral replication and liver fibrosis[24] and may induce cell apoptosis[70].

 To investigate the mechanism of the generation of pre-S deletion—that is, whether these pre-S deletion mutants are generated through RNA splicing or sporadic events—the splice donor and acceptor sites of the pre-S region have been searched, and only one type of pre-S1 deletion mutant was determined to have splice donor (nt 3018) and acceptor (nt 3202)(the nucleotide positions are based on genotypes B and C) site-specific sequences at the deletion boundaries. This suggests that these pre-S1 deletion mutants (spPS1) were derived from spliced pgRNA (Figure 2B)[25]. The splice donor site was at the existing position 3018 (nucleotide position 2985 based on HBV/D), whereas the splice acceptor site at position 3202 (nucleotide position 3169 based on HBV/D) was new (Table 1). Splice mapping revealed that the splice donor and splice acceptor residues critical for spPS1 were conserved across HBV genotypes A–H (Table 1). This phenomenon explains why this splice variant is frequently found during persistent viral infection[25,53,71-76].

The molecular characteristics of the novel splice variant spPS1 are mostly unknown. The splicing event of spPS1 results in a 183-nucleotide deletion in the C-terminal half of the pre-S1 region, complete deletion of two functional sites (the S promoter and site for dual topology), partial deletion of the NBS, and generation of a spliced L protein (spL, deletion of 61 amino acids, aa 58-118) (Figure 2B). S promoter deletion should lead to a reduction in S RNAs (consequently resulting in a low level or absence of M and S proteins) and an increase in pre-S1 RNAs (consequently resulting in relative overexpression of the spL surface protein). The removal of sites for dual topology and nucleocapsid binding in the spPS1 variant leads to uniform (e-Pre-S) conformation of spL proteins (Figure 4B) and decreased secretion of HBsAg and viral particles. Our *in vitro* study revealed that spPS1 (previously named dpS1) has a defect in S RNA transcription and secretion of envelope proteins[52]. Other studies have also demonstrated that spPS1 possesses a defect in secretion of envelope proteins, viral packaging, and subsequent virion secretion[53,71,72]. Western blot analysis showed that intracellular spL proteins exhibited a heterogeneous pattern, and additional spL proteins with a higher molecular weight were detected[52]. Immunofluorescence staining revealed that spL proteins were accumulated within the ER and displayed a granular staining pattern[52].

The clinical significance of the spPS1variant remains largely unknown. This variant has been found in an occult HBV–infected child[73] and numerous chronically HBV-infected patients worldwide, and it has frequently been found in the sera of individuals with CH and cirrhosis[53,71,72,74-76]. Clinical follow-up studies conducted over a period of 10-14 years indicate that after this variant occurs, acute exacerbation of CHB occurs, which is followed by the development of liver fibrosis[71,72]. A study demonstrated that the prevalence of spPS1 was higher in CH patients (7 of 55, 12.7% *vs* 1 of 55, 1.8%; *P* = 0.06) and LC patients (8 of 55, 14.5% *vs* 1 of 55, 1.8%; *P* = 0.032) than in ASCs[25]. Logistic regression analysis revealed that spPS1 variants were highly related to CH (*P* = 0.058) and significantly related to LC (*P* = 0.040). Thus, these clinical studies strongly suggest that the spPS1 variant could cause acute exacerbation of CHB, liver inﬂammation, and fibrosis.

***C-terminus S mutant and genotype***

The C-terminus S domain (aa 179-226) is hydrophobic and assumed to be inserted in the ER membrane (Figure 4A). This domain is involved in mediating the transit of envelope glycoproteins across the endoplasmic reticulum[77]. Mutations in this domain can result in a stable, glycosylated, but nonsecreted chain, thus affecting the biogenesis and secretion of subviral particles[78]. Two C-terminus S mutations were found and significantly correlated with HCC: P203Q (4/23, 17.4% in HCC *vs* 1/105, 1.0% in non-HCC, *P* = 0.004); S210R (8/23, 34.8% in HCC *vs* 4/105, 3.8% in non-HCC, *P* < 0.001); P203Q + S210R (4/23, 17.4% in HCC *vs* 0/110, 0% in non-HCC, *P* = 0.001)[79]. *In vitro* experiments revealed that P203Q, S210R, and P203Q+S210R significantly reduced the ratio of secreted and intracellular HBsAg compared with WT at each time point analyzed (*P* < 0.05); P203Q and P203Q+S210R increased the percentage of cells in S-phase compared with WT (P203Q: 26% ± 13%; P203Q+S210R: 29% ± 14%; WT: 18% ± 9%, *P* < 0.01); S210R increased the percentage of cells in the G2/M-phase (33% ± 6% for S210R *vs* 26% ± 8% for WT, *P* < 0.001)[79]. These results show that these two C-terminus S mutations, P203Q and S210R, hamper HBsAg secretion and are associated with increased cellular proliferation, supporting their involvement in HCC development. This study was conducted in patients infected with HBV/D or HBV/A. Additional studies are required to evaluate whether these mutations exist in other HBV genotypes and whether the conclusions of previous studies are valid.

***Pre-S/S nonsense mutation***

Pre-S nonsense mutations were also found in patients with progressive liver diseases[6]; the pathogenic impacts of these naturally occurring mutants remain unknown. Such pre-S nonsense mutations result in the occurrence of pre-S stop codon mutants and the synthesis of C terminally truncated M (MHBst) and L (LHBst) proteins. Studies have reported that MHBst and LHBst function as a transcriptional activator and result in an increased hepatocyte proliferation rate[78,80,81]. Results from experiments conducted on transgenic mice and hepatoma cell cultures have revealed that MHBst proteins retained in the ER can trigger a PKC dependent activation of the c-Raf-1/Erk2 signaling cascade, which leads to the induction of AP-1 and nuclear factor-kappa B (NF-κB) transcription factors as well as to enhanced proliferative activity of hepatocytes[82,83]. By contrast, Yeh *et al*[18] demonstrated that five patients who carried stop codons (nonsense mutation) in the pre-S region had a more favorable disease-free prognosis following multivariate analysis.

S nonsense mutations can arise as result of mutations in the P ORF that are generally caused by exposure to antivirals, a phenomenon commonly called antiviral drug–associated S gene mutations[84-87]. These mutations can cause the occurrence of S stop codon mutants and the synthesis of C terminally truncated L, M, and S proteins. For example, the HBV mutation that encodes rtA181T is selected in the viral polymerase during antiviral drug therapy and can also encode a stop codon in the overlapping S gene at amino acid 172 (sW172\*), resulting in truncation of the last 55 amino acids of the C-terminal hydrophobic region of the S domain. *In vitro* study revealed that the sW172\* variant had a secretory defect and exerted a dominant negative effect on WT HBV virion secretion[85]. In addition, sW172\* transgenic mice developed HCC in an *in vivo* study[86]. Other S nonsense mutants such as sC69\*, sL95\*, sW182\*, and sL216\* were identified in HCC tumors[87,88]. Functional studies of sL95\*, sW182\*, and sL216\* demonstrated that they had higher cell proliferation activities and transformation abilities than WT S, especially sW182\*[87]. The sW182\* mutant in HBV/C was also shown to be associated with liver cirrhosis[89].

***Possible pathogenesis of pre-S/S variants***

On the basis of the previous studies investigating pre-S/S variants, a model to explain the occurrences of the pre-S/S variants and the possible role of these mutants in progressive liver diseases is proposed (Figure 5). After persistent HBV infection, under the pressure of immune responses and antiviral drugs, immune epitope deletion and mutation occur along with drug-resistant mutants. Different pre-S deletion and pre-S/S mutants use different routes to cause liver diseases. Most cause the intracellular retention of HBV envelope proteins and induction of ER stress, resulting in liver diseases. Based on the region mutated, at least six pre-S/S variants occurred.

Type I — pre-S deletion in the N-terminal pre-S1 region causes deletion of the start codon of L proteins. *In vitro* study demonstrated that the L-start codon deletion mutant resulted in the absence of L proteins and increased levels of intracellular viral mRNA and extracellular HBsAg[56]. The accumulated intracellular viral mRNA might activate the intracellular toll-like receptors, leading to the subsequent activation of NF-κB pathways, chronic inﬂammation, and carcinogenesis[56].

Type II — pre-S deletion in the C-terminal half of the pre-S1 region can be separated into two groups characterized by S promoter: (II-a) S promoter deletion variants and (II-b) non-S promoter deletion variants (Figure 2B). (II-a) The S promoter deletion variants that cannot transcribe S RNA efficiently result in no synthesis or reduction of the M and S proteins. Because the L protein cannot be secreted from cells efficiently when expressed by itself, it must complex with the S and M proteins to form subviral particles or mature virions, bud from intracellular post-ER pre-Golgi membranes, and be released from the cell through secretion[41]. A low level or absence of M and S proteins results in the accumulation of mutant L proteins in the ER. *In vitro* studies have revealed severe intracellular retention of mutant L proteins in S promoter–variant transfected cells[52-56]. (II-b) The non-S promoter deletion variants can transcribe S RNA efficiently and synthesize the M and S proteins (Figure 2B), but the mutant L proteins may delete the T site to maintain a e-Pre-S form (Figure 4B, Topology variant) or delete the NBS site (Figure 4B, Nucleocapsid variant), leading to inefficient assembly of the nucleocapsid, viral immaturities, and mild intracellular retention[52].

Type III — N-terminus pre-S2 deletion mutants can also be separated into two groups by the start codon of the M protein: (III-a) non-M start codon deletion variants; (III-b) M start codon deletion variants; and (III-a) non-M start codon variants with internal deletion of M proteins but the mutant L proteins may lose the NBS or VS site, resulting in viral immaturities, and slight intracellular retention[52,90]. (III-b) The M start codon deletion variants with no M proteins change the ratio of mutant L, M, and S proteins, and lead to intracellular retention of mutant L proteins[52]. Because the M start codon is located in the NBS and VS sites, these variants may produce mutant L proteins such as the nucleocapsid variant that cannot assemble the nucleocapsid efficiently, leading to viral immaturities and slight intracellular retention of HBsAg[52].

Type IV — spPS1 variants are generated through RNA splicing of HBV pregenomic RNA. The splicing event results in a 183-nucleotide deletion in the C-terminal half of the pre-S1 region, complete deletion of two functional sites (the S promoter and T sites), partial deletion of the NBS site, and generation of spL (Figures 2B and 4B). S promoter deletion leads to absence of M and S proteins and severe intracellular retention of spL proteins. T-site deletion results in uniform conformation of spL proteins (Figure 4B) and loss of i-Pre-S form for capsid envelopment, which causes viral immaturities and intracellular retention of spL proteins.

Type V—C-terminus S mutants influence protein folding in the ER membrane, thus impairing HBsAg release, resulting in its accumulation in specific intracellular compartments (presumably represented by the ER and Golgi apparatus) and in turn contributing to cell proliferation[79]. An *in vitro* study revealed that C-terminus S mutants can also activate the proliferation control[79].

Type VI — pre-S/S nonsense mutations can be separated into two groups: (VI-a) Pre-S nonsense mutation and (VI-b) S nonsense mutation. (VI-a) Pre-S nonsense mutations can create C’ truncated L and M proteins, leading to transactivation of proliferation control and causing liver diseases[78,80,81]. (VI-b) S nonsense mutation can create C’ truncated L, M, and S proteins. *In vitro* study revealed that a stop codon in the C-terminal hydrophobic region of the S region results in truncated envelope proteins that are less glycosylated and are defective in secretion of viral particles, causing intracellular retention of envelope proteins and liver diseases[85]. *In vitro* and *in vivo* studies have also demonstrated that these S stop codon mutants have higher cell proliferation activity[86,87].

**CONCLUSION**

Naturally occurring pre-S/S variants are frequently found in chronically HBV-infected patients and have been identiﬁed as inﬂuencing liver disease progression. From a review of relevant studies, pre-S/S variants should be routinely determined in HBV carriers to help identify those who may be at a higher risk of a less favorable liver disease progression. In the future, further studies are required exploring the molecular mechanisms of the pre-S/S variants involved in the pathogenesis of each stage of liver disease.

**REFERENCES**

1 **Seeger C**, Mason WS. Hepatitis B virus biology. *Microbiol Mol Biol Rev* 2000; **64**: 51-68 [PMID: 10704474]

2 **Kao JH**, Chen DS. Global control of hepatitis B virus infection. *Lancet Infect Dis* 2002; **2**: 395-403 [PMID: 12127351]

3 **Liu CJ**, Kao JH. Global perspective on the natural history of chronic hepatitis B: role of hepatitis B virus genotypes A to J. *Semin Liver Dis* 2013; **33**: 97-102 [PMID: 23749665 DOI: 10.1055/s-0033-1345716]

4 **Lin CL**, Kao JH. The clinical implications of hepatitis B virus genotype: Recent advances. *J Gastroenterol Hepatol* 2011; **26 Suppl 1**: 123-130 [PMID: 21199523 DOI: 10.1111/j.1440-1746.2010.06541]

5 **Sugauchi F**, Ohno T, Orito E, Sakugawa H, Ichida T, Komatsu M, Kuramitsu T, Ueda R, Miyakawa Y, Mizokami M. Influence of hepatitis B virus genotypes on the development of preS deletions and advanced liver disease. *J Med Virol* 2003; **70**: 537-544 [PMID: 12794715]

6 **Chen BF**, Liu CJ, Jow GM, Chen PJ, Kao JH, Chen DS. High prevalence and mapping of pre-S deletion in hepatitis B virus carriers with progressive liver diseases. *Gastroenterology* 2006; **130**: 1153-1168 [PMID: 16618410]

7 **Okamoto H**, Imai M, Kametani M, Nakamura T, Mayumi M. Genomic heterogeneity of hepatitis B virus in a 54-year-old woman who contracted the infection through materno-fetal transmission. *Jpn J Exp Med* 1987; **57**: 231-236 [PMID: 3430800]

8 **Günther S**, Fischer L, Pult I, Sterneck M, Will H. Naturally occurring variants of hepatitis B virus. *Adv Virus Res* 1999; **52**: 25-137 [PMID: 10384235]

9 **Locarnini S**, McMillan J, Bartholomeusz A. The hepatitis B virus and common mutants. *Semin Liver Dis* 2003; **23**: 5-20 [PMID: 12616447 DOI: 10.1055/s-2003-37587]

10 **Hunt CM**, McGill JM, Allen MI, Condreay LD. Clinical relevance of hepatitis B viral mutations. *Hepatology* 2000; **31**: 1037-1044 [PMID: 10796877 DOI: 10.1053/he.2000.6709]

11 **Pollicino T**, Zanetti AR, Cacciola I, Petit MA, Smedile A, Campo S, Sagliocca L, Pasquali M, Tanzi E, Longo G, Raimondo G. Pre-S2 defective hepatitis B virus infection in patients with fulminant hepatitis. *Hepatology* 1997; **26**: 495-499 [PMID: 9252165 DOI: 10.1002/hep.510260235]

12 **Fan YF**, Lu CC, Chen WC, Yao WJ, Wang HC, Chang TT, Lei HY, Shiau AL, Su IJ. Prevalence and significance of hepatitis B virus (HBV) pre-S mutants in serum and liver at different replicative stages of chronic HBV infection. *Hepatology* 2001; **33**: 277-286 [PMID: 11124846 DOI: 10.1053/jhep.2001.21163]

13 **Chen CH**, Hung CH, Lee CM, Hu TH, Wang JH, Wang JC, Lu SN, Changchien CS. Pre-S deletion and complex mutations of hepatitis B virus related to advanced liver disease in HBeAg-negative patients. *Gastroenterology* 2007; **133**: 1466-1474 [PMID: 17915220 DOI: 10.1053/j.gastro.2007.09.002]

14 **Yeung P**, Wong DK, Lai CL, Fung J, Seto WK, Yuen MF. Association of hepatitis B virus pre-S deletions with the development of hepatocellular carcinoma in chronic hepatitis B. *J Infect Dis* 2011; **203**: 646-654 [PMID: 21227916 DOI: 10.1093/infdis/jiq096]

15 **Kao JH**, Chen PJ, Lai MY, Chen DS. Basal core promoter mutations of hepatitis B virus increase the risk of hepatocellular carcinoma in hepatitis B carriers. *Gastroenterology* 2003; **124**: 327-334 [PMID: 12557138 DOI: 10.1053/gast.2003.50053]

16 **Liu CJ**, Chen BF, Chen PJ, Lai MY, Huang WL, Kao JH, Chen DS. Role of hepatitis B viral load and basal core promoter mutation in hepatocellular carcinoma in hepatitis B carriers. *J Infect Dis* 2006; **193**: 1258-1265 [PMID: 16586363 DOI: 10.1086/502978]

17 **Chen CJ**, Yang HI, Su J, Jen CL, You SL, Lu SN, Huang GT, Iloeje UH; REVEAL-HBV Study Group. Risk of hepatocellular carcinoma across a biological gradient of serum hepatitis B virus DNA level. *JAMA* 2006; **295**: 65-73 [PMID: 16391218 DOI: 10.1001/jama.295.1.65]

18 **Yeh CT**, So M, Ng J, Yang HW, Chang ML, Lai MW, Chen TC, Lin CY, Yeh TS, Lee WC. Hepatitis B virus-DNA level and basal core promoter A1762T/G1764A mutation in liver tissue independently predict postoperative survival in hepatocellular carcinoma. *Hepatology* 2010; **52**: 1922-1933 [PMID: 20814897 DOI: 10.1002/hep.23898]

19 **Lin CL**, Kao JH. Risk stratification for hepatitis B virus related hepatocellular carcinoma. *J Gastroenterol Hepatol* 2013; **28**: 10-17 [PMID: 23094699 DOI: 10.1111/jgh.12010]

20 **Liu S**, Zhang H, Gu C, Yin J, He Y, Xie J, Cao G. Associations between hepatitis B virus mutations and the risk of hepatocellular carcinoma: a meta-analysis. *J Natl Cancer Inst* 2009; **101**: 1066-1082 [PMID: 19574418 DOI: 10.1093/jnci/djp180]

21 **Zhang ZH**, Wu CC, Chen XW, Li X, Li J, Lu MJ. Genetic variation of hepatitis B virus and its significance for pathogenesis. *World J Gastroenterol* 2016; **22**: 126-144 [PMID: 26755865 DOI: 10.3748/wjg.v22.i1.126]

22 **Lee D**, Lyu H, Chung YH, Kim JA, Mathews P, Jaffee E, Zheng L, Yu E, Lee YJ, Ryu SH. Genomic change in hepatitis B virus associated with development of hepatocellular carcinoma. *World J Gastroenterol* 2016; **22**: 5393-5399 [PMID: 27340355 DOI: 10.3748/wjg.v22.i23.5393]

23 **Soussan P**, Pol J, Garreau F, Schneider V, Le Pendeven C, Nalpas B, Lacombe K, Bonnard P, Pol S, Kremsdorf D. Expression of defective hepatitis B virus particles derived from singly spliced RNA is related to liver disease. *J Infect Dis* 2008; **198**: 218-225 [PMID: 18532883 DOI: 10.1086/589623]

24 **Soussan P**, Tuveri R, Nalpas B, Garreau F, Zavala F, Masson A, Pol S, Brechot C, Kremsdorf D. The expression of hepatitis B spliced protein (HBSP) encoded by a spliced hepatitis B virus RNA is associated with viral replication and liver fibrosis. *J Hepatol* 2003; **38**: 343-348 [PMID: 12586301]

25 **Lin YM**, Chen BF. A putative hepatitis B virus splice variant associated with chronic hepatitis and liver cirrhosis. *Virology* 2017; **510**: 224-233 [PMID: 28750326 DOI: 10.1016/j.virol.2017.07.025]

26 **Pollicino T**, Cacciola I, Saffioti F, Raimondo G. Hepatitis B virus PreS/S gene variants: pathobiology and clinical implications. *J Hepatol* 2014; **61**: 408-417 [PMID: 24801416 DOI: 10.1016/j.jhep.2014.04.041]

27 **Su IJ**, Wang HC, Wu HC, Huang WY. Ground glass hepatocytes contain pre-S mutants and represent preneoplastic lesions in chronic hepatitis B virus infection. *J Gastroenterol Hepatol* 2008; **23**: 1169-1174 [PMID: 18505413 DOI: 10.1111/j.1440-1746.2008.05348.x]

28 **Neurath AR**, Kent SB, Strick N, Parker K. Identification and chemical synthesis of a host cell receptor binding site on hepatitis B virus. *Cell* 1986; **46**: 429-436 [PMID: 3015414]

29 **Pontisso P**, Ruvoletto MG, Gerlich WH, Heermann KH, Bardini R, Alberti A. Identification of an attachment site for human liver plasma membranes on hepatitis B virus particles. *Virology* 1989; **173**: 522-530 [PMID: 2480688]

30 **Le Seyec J**, Chouteau P, Cannie I, Guguen-Guillouzo C, Gripon P. Role of the pre-S2 domain of the large envelope protein in hepatitis B virus assembly and infectivity. *J Virol* 1998; **72**: 5573-5578 [PMID: 9621015]

31 **Poisson F**, Severac A, Hourioux C, Goudeau A, Roingeard P. Both pre-S1 and S domains of hepatitis B virus envelope proteins interact with the core particle. *Virology* 1997; **228**: 115-120 [PMID: 9024817]

32 **Bruss V**. A short linear sequence in the pre-S domain of the large hepatitis B virus envelope protein required for virion formation. *J Virol* 1997; **71**: 9350-9357 [PMID: 9371594]

33 **Bock CT**, Kubicka S, Manns MP, Trautwein C. Two control elements in the hepatitis B virus S-promoter are important for full promoter activity mediated by CCAAT-binding factor. *Hepatology* 1999; **29**: 1236-1247 [PMID: 10094970]

34 **Sobotta D**, Sominskaya I, Jansons J, Meisel H, Schmitt S, Heermann KH, Kaluza G, Pumpens P, Gerlich WH. Mapping of immunodominant B-cell epitopes and the human serum albumin-binding site in natural hepatitis B virus surface antigen of defined genosubtype. *J Gen Virol* 2000; **81**: 369-378 [PMID: 10644835]

35 **Milich DR**, Jones JE, McLachlan A, Bitter G, Moriarty A, Hughes JL. Importance of subtype in the immune response to the pre-S(2) region of the hepatitis B surface antigen. II. Synthetic Pre-S(2) immunogen. *J Immunol* 1990; **144**: 3544-3551 [PMID: 1691763]

36 **Park JH**, Cho EW, Lee YJ, Shin SY, Kim KL. Determination of the protective effects of neutralizing anti-hepatitis B virus (HBV) immunoglobulins by epitope mapping with recombinant HBV surface-antigen proteins. *Microbiol Immunol* 2000; **44**: 703-710 [PMID: 11021401]

37 **Kuroki K**, Floreani M, Mimms LT, Ganem D. Epitope mapping of the PreS1 domain of the hepatitis B virus large surface protein. *Virology* 1990; **176**: 620-624 [PMID: 1693249]

38 **Maeng CY**, Ryu CJ, Gripon P, Guguen-Guillouzo C, Hong HJ. Fine mapping of virus-neutralizing epitopes on hepatitis B virus PreS1. *Virology* 2000; **270**: 9-16 [PMID: 10772975]

39 **Ferrari C**, Cavalli A, Penna A, Valli A, Bertoletti A, Pedretti G, Pilli M, Vitali P, Neri TM, Giuberti T. Fine specificity of the human T-cell response to the hepatitis B virus preS1 antigen. *Gastroenterology* 1992; **103**: 255-263 [PMID: 1377142]

40 **Huang TJ**, Lu CC, Tsai JC, Yao WJ, Lu X, Lai MD, Liu HS, Shiau AL. Novel autoregulatory function of hepatitis B virus M protein on surface gene expression. *J Biol Chem* 2005; **280**: 27742-27754 [PMID: 15899887 DOI: 10.1074/jbc.M502209200.]

41 **Bruss V**. Hepatitis B virus morphogenesis. *World J Gastroenterol* 2007; **13**: 65-73 [PMID: 17206755 DOI: 10.3748/wjg.v13.i1.65]

42 **Lambert C**, Prange R. Chaperone action in the posttranslational topological reorientation of the hepatitis B virus large envelope protein: Implications for translocational regulation. *Proc Natl Acad Sci U S A* 2003; **100**: 5199-5204 [PMID: 12697898 DOI: 10.1073/pnas.0930813100]

43 **Löffler-Mary H**, Werr M, Prange R. Sequence-specific repression of cotranslational translocation of the hepatitis B virus envelope proteins coincides with binding of heat shock protein Hsc70. *Virology* 1997; **235**: 144-152 [PMID: 9300046 DOI: 10.1006/viro.1997.8689]

44 **Zhang AY**, Lai CL, Huang FY, Seto WK, Fung J, Wong DK, Yuen MF. Evolutionary Changes of Hepatitis B Virus Pre-S Mutations Prior to Development of Hepatocellular Carcinoma. *PLoS One* 2015; **10**: e0139478 [PMID: 26421619 DOI: 10.1371/journal.pone.0139478.]

45 **Huang HP**, Hsu HY, Chen CL, Ni YH, Wang HY, Tsuei DJ, Chiang CL, Tsai YC, Chen HL, Chang MH. Pre-S2 deletions of hepatitis B virus and hepatocellular carcinoma in children. *Pediatr Res* 2010; **67**: 90-94 [PMID: 19816238 DOI: 10.1203/PDR.0b013e3181c1b0b7]

46 **Abe K**, Thung SN, Wu HC, Tran TT, Le Hoang P, Truong KD, Inui A, Jang JJ, Su IJ. Pre-S2 deletion mutants of hepatitis B virus could have an important role in hepatocarcinogenesis in Asian children. *Cancer Sci* 2009; **100**: 2249-2254 [PMID: 19719772 DOI: 10.1111/j.1349-7006.2009.01309.x]Plos]

47 **Mun HS**, Lee SA, Jee Y, Kim H, Park JH, Song BC, Yoon JH, Kim YJ, Lee HS, Hyun JW, Hwang ES, Kook YH, Kim BJ. The prevalence of hepatitis B virus preS deletions occurring naturally in Korean patients infected chronically with genotype C. *J Med Virol* 2008; **80**: 1189-1194 [PMID: 18461612 DOI: 10.1002/jmv.21208]

48 **Ahn SH**, Yuen L, Han KH, Littlejohn M, Chang HY, Damerow H, Ayres A, Heo J, Locarnini S, Revill PA. Molecular and clinical characteristics of hepatitis B virus in Korea. *J Med Virol* 2010; **82**: 1126-1134 [PMID: 20513074 DOI: 10.1002/jmv.21844]

49 **Biswas A**, Panigrahi R, Banerjee A, Pal M, De BK, Chakrabarti S, Chakravarty R. Differential pattern of pre-S mutations/deletions and its association with hepatitis B virus genotypes in Eastern India. *Infect Genet Evol* 2012; **12**: 384-391 [PMID: 22266243 DOI: 10.1016/j.meegid.2012.01.007]

50 **Chen BF**. Different pre-S deletion patterns and their association with hepatitis B virus genotypes. *World J Gastroenterol* 2016; **22**: 8041-8049 [PMID: 27672298 DOI: 10.3748/wjg.v22.i35.8041.]

51 **Kao JH**, Liu CJ, Jow GM, Chen PJ, Chen DS, Chen BF. Fine mapping of hepatitis B virus pre-S deletion and its association with hepatocellular carcinoma. *Liver Int* 2012; **32**: 1373-1381 [PMID: 22676233]

52 **Lin CM**, Wang GM, Jow GM, Chen BF. Functional analysis of hepatitis B virus pre-s deletion variants associated with hepatocellular carcinoma. *J Biomed Sci* 2012; **19**: 17 [PMID: 22313590 DOI: 10.1186/1423-0127-19-17]

53 **Melegari M**, Scaglioni PP, Wands JR. The small envelope protein is required for secretion of a naturally occurring hepatitis B virus mutant with pre-S1 deleted. *J Virol* 1997; **71**: 5449-5454 [PMID: 9188617]

54 **Bock CT**, Tillmann HL, Maschek HJ, Manns MP, Trautwein C. A preS mutation isolated from a patient with chronic hepatitis B infection leads to virus retention and misassembly. *Gastroenterology* 1997; **113**: 1976-1982 [PMID: 9394738]

55 **Xu Z**, Yen TS. Intracellular retention of surface protein by a hepatitis B virus mutant that releases virion particles. *J Virol* 1996; **70**: 133-140 [PMID: 8523517]

56 **Kim BK**, Choi SH, Ahn SH, Chung AR, Park YK, Han KH, Kim S, Kim HS, Park JH, Kim KS, Lee HS, Cho YS, Kim KH, Ahn SH. Pre-S mutations of hepatitis B virus affect genome replication and expression of surface antigens. *J Gastroenterol Hepatol* 2014; **29**: 843-850 [PMID: 24783251]

57 **Yang JC**, Teng CF, Wu HC, Tsai HW, Chuang HC, Tsai TF, Hsu YH, Huang W, Wu LW, Su IJ. Enhanced expression of vascular endothelial growth factor-A in ground glass hepatocytes and its implication in hepatitis B virus hepatocarcinogenesis. *Hepatology* 2009; **49**: 1962-1971 [PMID: 19475690 DOI: 10.1002/hep.22889]

58 **Wang HC**, Chang WT, Chang WW, Wu HC, Huang W, Lei HY, Lai MD, Fausto N, Su IJ. Hepatitis B virus pre-S2 mutant upregulates cyclin A expression and induces nodular proliferation of hepatocytes. *Hepatology* 2005; **41**: 761-770 [PMID: 15726643 DOI: 10.1002/hep.20615]

59 **Wang LH**, Huang W, Lai MD, Su IJ. Aberrant cyclin A expression and centrosome overduplication induced by hepatitis B virus pre-S2 mutants and its implication in hepatocarcinogenesis. *Carcinogenesis* 2012; **33**: 466-472 [PMID: 22159224 DOI: 10.1093/carcin/bgr296]

60 **Hung JH**, Su IJ, Lei HY, Wang HC, Lin WC, Chang WT, Huang W, Chang WC, Chang YS, Chen CC, Lai MD. Endoplasmic reticulum stress stimulates the expression of cyclooxygenase-2 through activation of NF-kappaB and pp38 mitogen-activated protein kinase. *J Biol Chem* 2004; **279**: 46384-46392 [PMID: 15319438 DOI: 10.1074/jbc.M403568200]

61 **Hsieh YH**, Su IJ, Wang HC, Tsai JH, Huang YJ, Chang WW, Lai MD, Lei HY, Huang W. Hepatitis B virus pre-S2 mutant surface antigen induces degradation of cyclin-dependent kinase inhibitor p27Kip1 through c-Jun activation domain-binding protein 1. *Mol Cancer Res* 2007; **5**: 1063-1072 [PMID: 17951406 DOI: 10.1158/1541-7786.MCR-07-0098]

62 **Chen CH**, Changchien CS, Lee CM, Hung CH, Hu TH, Wang JH, Wang JC, Lu SN. Combined mutations in pre-s/surface and core promoter/precore regions of hepatitis B virus increase the risk of hepatocellular carcinoma: a case-control study. *J Infect Dis* 2008; **198**: 1634-1642 [PMID: 18939932 DOI: 10.1086/592990]

63 **Zhang AY**, Lai CL, Huang FY, Seto WK, Fung J, Wong DK, Yuen MF. Deep sequencing analysis of quasispecies in the HBV pre-S region and its association with hepatocellular carcinoma. *J Gastroenterol* 2017; **52**: 1064-1074 [PMID: 28353014 DOI: 10.1007/s00535-017-1334-1.]

64 **Zhang Q**, Cao G. Genotypes, mutations, and viral load of hepatitis B virus and the risk of hepatocellular carcinoma: HBV properties and hepatocarcinogenesis. *Hepat Mon* 2011; **11**: 86-91 [PMID: 22087123]

65 **Mun HS**, Lee SA, Kim H, Hwang ES, Kook YH, Kim BJ. Novel F141L pre-S2 mutation in hepatitis B virus increases the risk of hepatocellular carcinoma in patients with chronic genotype C infections. *J Virol* 2011; **85**: 123-132 [PMID: 20962085 DOI: 10.1128/JVI.01524-10.]

66 **Bayliss J**, Lim L, Thompson AJ, Desmond P, Angus P, Locarnini S, Revill PA. Hepatitis B virus splicing is enhanced prior to development of hepatocellular carcinoma. *J Hepatol* 2013; **59**: 1022-1028 [PMID: 23811301 DOI: 10.1016/j.jhep.2013.06.018]

67 **Günther S**, Sommer G, Iwanska A, Will H. Heterogeneity and common features of defective hepatitis B virus genomes derived from spliced pregenomic RNA. *Virology* 1997; **238**: 363-371 [PMID: 9400609 DOI: 10.1006/viro.1997.8863]

68 **Ma ZM**, Lin X, Wang YX, Tian XC, Xie YH, Wen YM. A double-spliced defective hepatitis B virus genome derived from hepatocellular carcinoma tissue enhanced replication of full-length virus. *J Med Virol* 2009; **81**: 230-237 [PMID: 19107969 DOI: 10.1002/jmv.21393]

69 **Abraham TM**, Lewellyn EB, Haines KM, Loeb DD. Characterization of the contribution of spliced RNAs of hepatitis B virus to DNA synthesis in transfected cultures of Huh7 and HepG2 cells. *Virology* 2008; **379**: 30-37 [PMID: 18657840 DOI: 10.1016/j.virol.2008.06.021]

70 **Soussan P**, Garreau F, Zylberberg H, Ferray C, Brechot C, Kremsdorf D. In vivo expression of a new hepatitis B virus protein encoded by a spliced RNA. *J Clin Invest* 2000; **105**: 55-60 [PMID: 10619861 DOI: 10.1172/JCI8098]

71 **Gerner P**, Schäfer HM, Prange R, Pravitt D, Wirth S. Functional analysis of a rare HBV deletion mutant in chronically infected children. *Pediatr Res* 2003; **53**: 891-897 [PMID: 12646724 DOI: 10.1203/01.PDR.0000064906.63939.72]

72 **Kajiya Y**, Hamasaki K, Nakata K, Nakagawa Y, Miyazoe S, Takeda Y, Ohkubo K, Ichikawa T, Nakao K, Kato Y, Eguchi K. Full-length sequence and functional analysis of hepatitis B virus genome in a virus carrier: a case report suggesting the impact of pre-S and core promoter mutations on the progression of the disease. *J Viral Hepat* 2002; **9**: 149-156 [PMID: 11876799 DOI: 10.1046/j.1365-2893.2002.00335.x]

73 **Mu SC**, Lin YM, Jow GM, Chen BF. Occult hepatitis B virus infection in hepatitis B vaccinated children in Taiwan. *J Hepatol* 2009; **50**: 264-272 [PMID: 19070923 DOI: 10.1016/j.jhep.2008.09.017]

74 **Nakajima E**, Minami M, Ochiya T, Kagawa K, Okanoue T. PreS1 deleted variants of hepatitis B virus in patients with chronic hepatitis. *J Hepatol* 1994; **20**: 329-335 [PMID: 7516949]

75 **Takayanagi M**, Kakumu S, Ishikawa T, Higashi Y, Yoshioka K, Wakita T. Comparison of envelope and precore/core variants of hepatitis B virus (HBV) during chronic HBV infection. *Virology* 1993; **196**: 138-145 [PMID: 8356791 DOI: 10.1006/viro.1993.1462]

76 **Gerner PR**, Friedt M, Oettinger R, Lausch E, Wirth S. The hepatitis B virus seroconversion to anti-HBe is frequently associated with HBV genotype changes and selection of preS2-defective particles in chronically infected children. *Virology* 1998; **245**: 163-172 [PMID: 9614877]

77 **Bruss V**, Ganem D. Mutational analysis of hepatitis B surface antigen particle assembly and secretion. *J Virol* 1991; **65**: 3813-3820 [PMID: 2041095]

78 **Kim HS**, Ryu CJ, Hong HJ. Hepatitis B virus preS1 functions as a transcriptional activation domain. *J Gen Virol* 1997; **78 (Pt 5)**: 1083-1086 [PMID: 9152426 DOI: 10.1099/0022-1317-78-5-1083]

79 **Salpini R**, Surdo M, Warner N, Cortese MF, Colledge D, Soppe S, Bellocchi MC, Armenia D, Carioti L, Continenza F, Di Carlo D, Saccomandi P, Mirabelli C, Pollicita M, Longo R, Romano S, Cappiello G, Spanò A, Trimoulet P, Fleury H, Vecchiet J, Iapadre N, Barlattani A, Bertoli A, Mari T, Pasquazzi C, Missale G, Sarrecchia C, Orecchini E, Michienzi A, Andreoni M, Francioso S, Angelico M, Verheyen J, Ceccherini-Silberstein F, Locarnini S, Perno CF, Svicher V. Novel HBsAg mutations correlate with hepatocellular carcinoma, hamper HBsAg secretion and promote cell proliferation in vitro. *Oncotarget* 2017; **8**: 15704-15715 [PMID: 28152517 DOI: 10.18632/oncotarget.14944]

80 **Hildt E**, Urban S, Hofschneider PH. Characterization of essential domains for the functionality of the MHBst transcriptional activator and identification of a minimal MHBst activator. *Oncogene* 1995; **11**: 2055-2066 [PMID: 7478525]

81 **Hildt E**, Saher G, Bruss V, Hofschneider PH. The hepatitis B virus large surface protein (LHBs) is a transcriptional activator. *Virology* 1996; **225**: 235-239 [PMID: 8918553 DOI: 10.1006/viro.1996.0594]

82 **Hildt E**, Munz B, Saher G, Reifenberg K, Hofschneider PH. The PreS2 activator MHBs(t) of hepatitis B virus activates c-raf-1/Erk2 signaling in transgenic mice. *EMBO J* 2002; **21**: 525-535 [PMID: 11847101 DOI: 10.1093/emboj/21.4.525]

83 **Lupberger J**, Hildt E. Hepatitis B virus-induced oncogenesis. *World J Gastroenterol* 2007; **13**: 74-81 [PMID: 17206756 DOI: 10.3748/wjg.v13.i1.74]

84 **Kamili S**, Sozzi V, Thompson G, Campbell K, Walker CM, Locarnini S, Krawczynski K. Efficacy of hepatitis B vaccine against antiviral drug-resistant hepatitis B virus mutants in the chimpanzee model. *Hepatology* 2009; **49**: 1483-1491 [PMID: 19274751 DOI: 10.1002/hep.22796]

85 **Warner N**, Locarnini S. 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* 2008; **48**: 88-98 [PMID: 18537180 DOI: 10.1002/hep.22295]

86 **Lai MW**, Liang KH, Lin WR, Huang YH, Huang SF, Chen TC, Yeh CT. Hepatocarcinogenesis in transgenic mice carrying hepatitis B virus pre-S/S gene with the sW172\* mutation. *Oncogenesis* 2016; **5**: e273 [PMID: 27918551 DOI: 10.1038/oncsis.2016.77.]

87 **Huang SF**, Chen YT, Lee WC, Chang IC, Chiu YT, Chang Y, Tu HC, Yuh CH, Matsuura I, Shih LY, Lai MW, Wu HD, Chen MF, Yeh CT. Identification of transforming hepatitis B virus S gene nonsense mutations derived from freely replicative viruses in hepatocellular carcinoma. *PLoS One* 2014; **9**: e89753 [PMID: 24587012 DOI: 10.1371/journal.pone.0089753]

88 **Shirvani-Dastgerdi E**, Winer BY, Celià-Terrassa T, Kang Y, Tabernero D, Yagmur E, Rodríguez-Frías F, Gregori J, Luedde T, Trautwein C, Ploss A, Tacke F. Selection of the highly replicative and partially multidrug resistant rtS78T HBV polymerase mutation during TDF-ETV combination therapy. *J Hepatol* 2017; **67**: 246-254 [PMID: 28392234 DOI: 10.1016/j.jhep.2017.03.027.PMID:]

89 **Lee SA**, Kim K, Kim H, Kim BJ. 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* 2012; **56**: 63-69 [PMID: 21827734 DOI: 10.1016/j.jhep.2011.06.028]

90 **Tai PC**, Suk FM, Gerlich WH, Neurath AR, Shih C. Hypermodification and immune escape of an internally deleted middle-envelope (M) protein of frequent and predominant hepatitis B virus variants. *Virology* 2002; **292**: 44-58 [PMID: 11878907]

**P-Reviewer:** El-Shabrawi MHF, Komatsu H, Rodriguez-Frias F, Takahashi T

**S-Editor:** Gong ZM **L-Editor:** **E-Editor:**

**Specialty type:** Gastroenterology and hepatology

**Country of origin:** Taiwan

**Peer-review report classification**

Grade A (Excellent): A, A

Grade B (Very good): B, B

Grade C (Good): 0

Grade D (Fair): 0

Grade E (Poor): 0

**Table1 Putative 5’ splice donor and 3’ splice acceptor sites in hepatitis B virus used to generate the splice variant spPS1**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Genotype** | **Position (nt)** | **type** | **Potential splice donor site** | **Position (nt)** | **type** | **Potential splice acceptor site** | **ref.** |
| A | 3024/3025 | Donor | CAG/gtagga | 3207/3208 | Acceptor | tcatcctcag/GC |  |
| B | 3018/3019 | Donor | AAG/gtggga | 3201/3202 | Acceptor | tcatcctcag/GC | [25,73] |
| C | 3018/3019 | Donor | CAG/gtagga | 3201/3202 | Acceptor | tcatcctcag/GC | [25,72,74] |
| D | 2985/2986 | Donor | AAG/gtagga | 3168/3169 | Acceptor | tcatcctcag/GC | [53,71,76] |
| E | 3015/3016 | Donor | AAG/gtagga | 3198/3199 | Acceptor | tcatcctcag/GC |  |
| F | 3018/3019 | Donor | AAG/gtagga | 3201/3202 | Acceptor | acatcctcag/GC |  |
| G | 3051/3052 | Donor | AAG/gtagga | 3234/3235 | Acceptor | tcatcctcag/GC |  |
| H | 3018/3019 | Donor | AAG/gtagga | 3201/3202 | Acceptor | acatccacag/GC |  |



**Figure 1 Genome structure and organization of hepatitis B virus**. The relaxed-circular DNA genome of HBV with a complete minus strand and incomplete plus strand is shown (inner circle), along with the four main open reading frames (ORFs): pre-S/S; precore/core (pC/C); Pol, including four domains: TP, SP, reverse transcriptase (RT), and RNase H (RH); and X. The minus (−) and plus (+) DNA strands are marked. The HBV Pol and capped mRNA oligomer at the 5' end of the (−) and (+) strands as well as the DR-1 and DR-2 are illustrated. The space between the DR-1 and DR-2 is the "cohesive overlap region." The (+) strand is typically incomplete.

A



B



**Figure 2** **Gene expression of the pre-S/S gene in (A) wild-type hepatitis B virus and (B) pre-S/S variants: non-S promoter, S promoter, and spPS1.**



**Figure 3 Immune epitopes and functional domains within the hepatitis B virus pre-S region.** The pre-S region consists of the pre-S1 and pre-S2 regions. The pre-S1 region contains 119 amino acids in HBV genotypes B or C and is further divided into two parts: the N half (amino acids 1–57) and C half (amino acids 58–119). The pre-S2 region contains 55 amino acids. The pre-S domain contains many B- or T-epitopes and exerts multiple functions, as illustrated. The N-half of pre-S1 contains a hepatocyte binding site essential for infection. The C-half of pre-S1 contains a heat-shock protein 70 (Hsc70) binding site and cytosolic anchorage determinant (CAD) vital for dual topology of L proteins as well as a nucleocapsid binding site (NBS) for virion morphogenesis. The C-half of pre-S1 also contains an S-promoter and CCAAT binding factor (CBF) binding site necessary for expression of the S gene. The pre-S2 region has a polymerized human serum albumin (pHSA) binging site and viral secretion (VS) site. Black triangle, myristylation at second amino acid; white triangle, N-link glycosylation at N-4 of the M protein; gray triangle, O-link glycosylation at T-37 of the M protein. B-epitopes: pS1-B1, pS1-B2, pS2-B1, pS2-B2, and pS2-B3. T-epitopes: pS1-T1 and pS2-T1. B- and T-epitope: pS1-BT.

A



B



**Figure 4** **Topology of wild-type small (S), medium (M), and large (L) envelope proteins. The predicted four membrane-spanning segments (indicated by rectangular boxes) of S project their N and C termini into the endoplasmic reticulum lumen (A).** The M proteins exhibit a topology similar to the S proteins with their N-terminal pre-S2 domain protruding into the endoplasmic reticulum (ER) lumen, whereas the L proteins display a dual topology. Upon cotranslational membrane integration, the pre-S domains of L proteins are initially located on the cytosolic side of the ER membrane (i-Pre-S); they are controlled by the dual topology site (indicated by an oval). During maturation (marked by the arrow), nearly half of mature L-protein molecules posttranslationally translocate their pre-S region to the luminal space (e-Pre-S). The nucleocapsid (N) binding sites in the pre-S and S region are indicated by the white curved box. (B) The L-protein topology of pre-S/S variants. The nucleocapsid variant demonstrates a dual topology, and the topology variants and spPS1 variants display a uniform topology. The broken line indicates deletion, and “a” indicates “a” determinant.



**Figure 5 Proposed model for the generation of pre-S/S variants and their possible roles in liver damage and carcinogenesis.**