

## RESEARCH ARTICLE

# HBV preS deletion mapping using deep sequencing demonstrates a unique association with viral markers

Yuichiro Suzuki<sup>1</sup>, Shinya Maekawa<sup>1\*</sup>, Nobutoshi Komatsu<sup>1</sup>, Mitsuki Sato<sup>1</sup>, Akihisa Tatsumi<sup>1</sup>, Mika Miura<sup>1</sup>, Shuya Matsuda<sup>1</sup>, Masaru Muraoka<sup>1</sup>, Natsuko Nakakuki<sup>1</sup>, Fumitake Amemiya<sup>1</sup>, Shinichi Takano<sup>1</sup>, Mitsuharu Fukasawa<sup>1</sup>, Yasuhiro Nakayama<sup>1</sup>, Tatsuya Yamaguchi<sup>1</sup>, Taisuke Inoue<sup>1</sup>, Tadashi Sato<sup>1</sup>, Minoru Sakamoto<sup>1</sup>, Atsuya Yamashita<sup>2</sup>, Kohji Moriishi<sup>2</sup>, Nobuyuki Enomoto<sup>1</sup>

**1** First Department of Internal Medicine, Faculty of Medicine, University of Yamanashi, Chuo, Yamanashi, Japan, **2** Department of Microbiology, University of Yamanashi, Chuo, Yamanashi, Japan

\* [maekawa@yamanashi.ac.jp](mailto:maekawa@yamanashi.ac.jp)



## OPEN ACCESS

**Citation:** Suzuki Y, Maekawa S, Komatsu N, Sato M, Tatsumi A, Miura M, et al. (2019) HBV preS deletion mapping using deep sequencing demonstrates a unique association with viral markers. PLoS ONE 14(2): e0212559. <https://doi.org/10.1371/journal.pone.0212559>

**Editor:** Jason Blackard, University of Cincinnati College of Medicine, UNITED STATES

**Received:** July 30, 2018

**Accepted:** February 5, 2019

**Published:** February 22, 2019

**Copyright:** This is an open access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the [Creative Commons CC0](https://creativecommons.org/licenses/by/4.0/) public domain dedication.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** This study was partly supported by Research Program on Hepatitis from Japanese Agency for Medical Research and Development (AMED) Japan (Grant Number 16fk0210301h0003, 16fk0210102h0001, and 16fk0210308h0003 to NE), and partly by JSPS KAKENHI (Grant Number 16K09349). This study was also supported partly from Japanese Agency

## Abstract

### Aim

Deletions are observed frequently in the preS1/S2 region of hepatitis B virus (HBV) genome, in association with liver disease advancement. However, the most significant preS1/S2 region and its influences on viral markers are unclear.

### Methods

The preS1/S2 HBV regions of 90 patients without antiviral therapy were subjected to deep sequencing and deleted regions influencing viral markers were investigated.

### Results

From the deletion frequency analysis in each patient, deletions were observed most frequently in the preS2 codon 132–141 region. When the patients were divided into three groups (0–0.1%: n = 27, 0.1%–10%: n = 34, 10–100%: n = 29), based on the deletion frequency, FIB-4 (p < 0.01), HBV DNA (p < 0.01), HBcrAg (p < 0.01) and preS1/S2 start codon mutations (p < 0.01, both) were significantly associated with the deletion. When clinical and viral markers were investigated by multivariate analysis for their association with the deletion, FIB-4 (p < 0.05), HBcrAg (p < 0.05), and preS1 start codon mutation (p < 0.01) were extracted as independent variables. When the influence of the preS codon 132–141 deletions on HBsAg and HBcrAg, relative to HBV DNA, was investigated, the HBsAg/HBV DNA ratio was lower (0–10% vs. 10%–100%, p < 0.05), while the HBcrAg/HBV DNA ratio was higher (0–0.1% vs. 10%–100%, p < 0.05) in the presence of the preS codon 132–141 deletions.

### Conclusion

The preS codon.132–141 deletions have a significant influence on the clinical characteristics and viral markers, even when present as a minor population. Importantly, the

for Medical Research and Development (AMED) Japan (Grant Number 17fk0310103j0001, 17fk0310113h0001 and 17fk0210202h0002 to NE). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

preS codon 132–141 deletions have a clear influence on the viral life cycle and pathogenesis.

## Introduction

Hepatitis B virus (HBV) chronically infects more than 257 million people worldwide and increases the risk of these individuals developing liver cirrhosis, hepatic decompensation and hepatocellular carcinoma (HCC) over the long course of the disease [1]. Recent advances in the development of nucleoside and nucleotide analogues (NAs) have made it possible to decrease hepatitis activity and to suppress serum hepatitis B virus DNA (HBV DNA) dramatically. However, it is also acknowledged that HCC may develop in a substantial number of patients, even after the introduction of these NAs, while prediction of those patients who will develop liver disease after NA introduction is difficult. Consequently, appropriate biomarkers that predict disease development are needed urgently. HBV markers, such as genomic sequences and viral proteins, are candidates for such biomarkers but the precise roles of these viral markers for disease advancement are not fully understood.

The preS region of the HBV genome comprises preS1 and preS2 and it is known that various mutations are often found there, along with liver disease advancement, and that deletions are the most frequent [2]. These mutations are considered to occur as a result of viral escape from the host's immune response, because the region contains B/T-cell epitopes [3–9]. It also has been reported that the preS mutations might influence the serum hepatitis B surface antigen (HBsAg) titer, because the preS region plays a role in HBsAg secretion from hepatocytes [10]. Considering this background, quantification of the preS mutations might improve our understanding of the mechanism of liver disease progression. On the other hand, it is not yet known which preS mutant is most important and how the contribution of the preS mutant to the viral quasispecies affects liver disease progression.

Recently, serum HBsAg quantification became possible and is considered an important viral marker, reflecting intrahepatic hepatitis B virus cccDNA (HBV cccDNA) [11] and, therefore, decreasing or even eliminating serum HBsAg is considered to be and has been proposed as the ultimate goal of anti-HBV therapy. More recently, the serum hepatitis B core-related antigen (HBcrAg) titer, a test developed in Japan to quantify the combined titer of serum hepatitis B core antigen (HBcAg), hepatitis B e antigen (HBeAg) and p22cr antigen (p22crAg) [12, 13], was also reported as an additional marker reflecting intrahepatic HBV cccDNA [14]. Because the presence of preS mutations could affect the serum HBsAg titer, as stated above, the interrelationship among the quasispecies state of preS mutants, HBsAg, HBcrAg and disease advancement is considered rather complicated. However, determining the quantitative interrelationships among these factors might advance our understanding of the pathogenesis of HBV-induced liver disease.

In this study, deep sequencing analysis of preS region was carried out to determine the most relevant preS deletion mutant associated with the development of liver fibrosis in chronic HBV patients and to disclose how the determined preS deletion affects the clinical characteristics, as well as viral markers.

## Results

### Clinical characteristics of the patients

The clinical backgrounds and viral markers of the 90 patients, including 29 inactive carriers, 28 with chronic hepatitis and 33 with cirrhosis, are shown in [Table 1](#).

**Table 1. Background of the patients.**

Factor	Inactive carrier (n = 29)	Chronic hepatitis (n = 28)	Cirrhosis (n = 33)
<b>Clinical parameters</b>			
Age (> 65 y/o), n(%)	5 (17%)	2 (8.3%)	9 (27%)
Gender (male)	13 (45%)	14 (50%)	25 (76%)
Platelets (<15 x10 <sup>4</sup> /mm <sup>3</sup> ), n(%)	4 (14%)	15 (50%)	28 (85%)
Alb (< 4.0 g/dl), n(%)	2 (6.8%)	6 (20%)	22 (67%)
ALT (>31 IU/l), n(%)	3 (10%)	24 (80%)	20 (61%)
AFP (≥5ng/ml), n(%)	2 (6.8%)	8 (27%)	21 (64%)
FIB-4* >1.45	11 (38%)	20 (71%)	33 (100%)
HCC, n(%)	0 (0%)	1 (3.5%)	14 (42%)
<b>Viral markers</b>			
HBeAg (+), n(%)	0 (0%)	10 (33%)	13 (39%)
HBsAg (< 3log IU/ml), n(%)	16 (55%)	12 (42%)	16 (48%)
HBV DNA (>4.0 log IU/ml), n (%)	0 (0%)	16 (53%)	17 (52%)
HBcrAg (log IU/ml), n(%)	5 (17%)	28 (93%)	30 (91%)
Genotype A/B/C, n(%)	1 (3.4%)/ 9 (31%)/ 19 (65.5%)	1 (3.6%)/ 0 (0%)/ 27 (96.4%)	0 (0%)/ 3 (9.1%)/ 30 (90.9%)

Alb, Albumin; ALT, alamine amino transferase; AFP, alfa-fetoprotein; FIB-4, Fibrosis-4; HCC, hepatocellular carcinoma; HBeAg, hepatitis B e antigen; HBsAg, hepatitis B surface antigen; HBV DNA, hepatitis B virus DNA; HBcrAg, hepatitis B core-related antigen.

<https://doi.org/10.1371/journal.pone.0212559.t001>

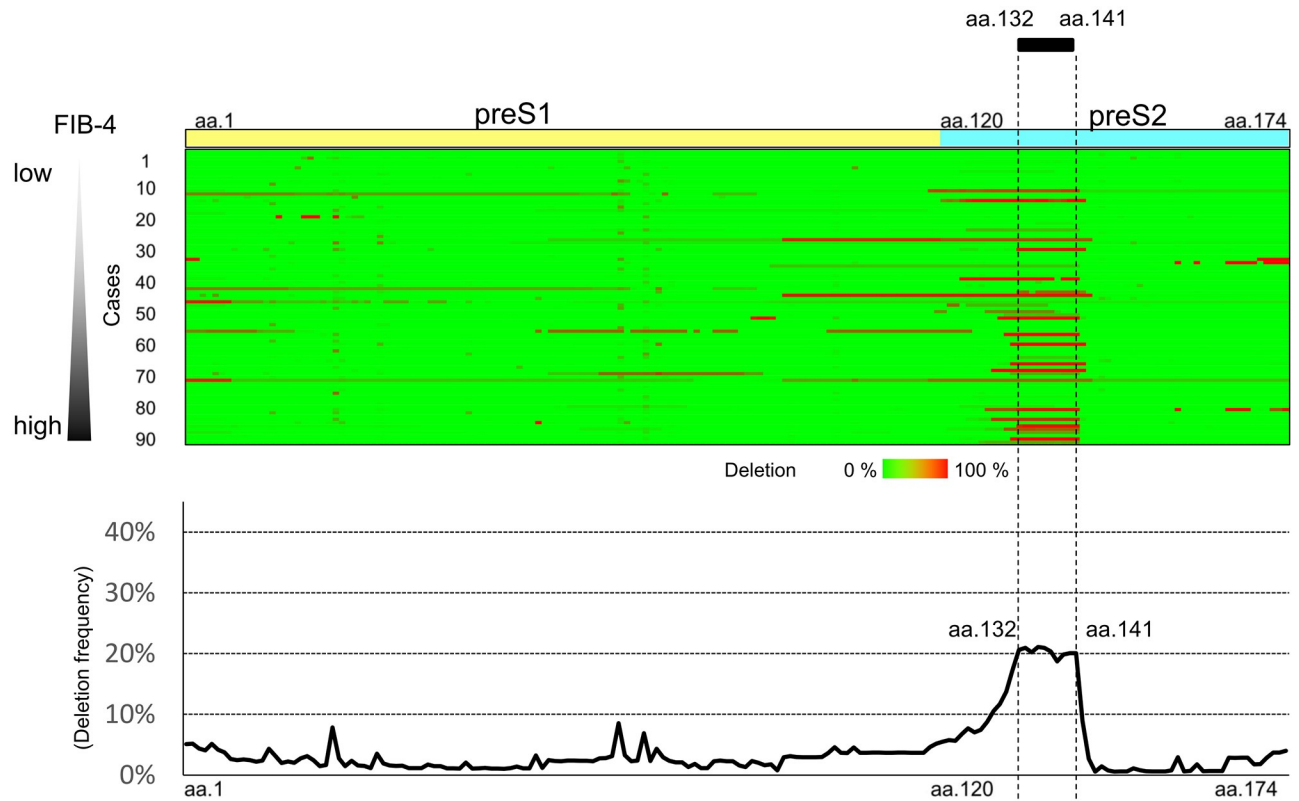
There were 52 males (58%), and 17% (15/90) were with HCC at the time of enrollment. As to viral markers, 26% (23/90) were HBeAg positive. Median values of HBsAg, HBV DNA, and HBcrAg were 3.0 log IU/ml, 5.5 log IU/ml and 4.2 log IU/ml, respectively. Eighty-four percent (76/90) of the patients were infected with genotype C HBV, while the others were infected with genotype A or B. Male sex, low platelets, low albumin, high alpha fetoprotein (AFP, a tumor marker for HCC), high Fibrosis-4 (FIB-4, an indicator of liver fibrosis) index and HCC history were observed more frequently in the liver cirrhosis group than the inactive carrier and chronic hepatitis groups (Table 1). Moreover, HBeAg positivity, high serum HBV DNA, high HBcrAg and genotype C were also more frequent in the cirrhosis group.

### Mapping of the preS region associated with platelet counts

Deep sequencing of the region of the HBV genome encoding the 174 preS amino acids was carried out. For each patient, the presence of amino acid deletions equal to or above the 0.1% cut-off, determined by the control plasmid experiment, was analyzed and the deleted amino acids were mapped for all 90 patients. In Fig 1 (upper panel), the deleted regions are colored red, while the region without deletion is colored green. In this figure, the preS region data of the 90 patients are arranged in the order of FIB-4 values. In Fig 1 (lower panel), the average deletion frequency of preS region per patient is shown. Most of the deletion were in-frame and were without frameshift or stop codon. The region of the genome encoding preS2 codon 132 to 141 was deleted most frequently.

### Deletion frequency in the preS region and its association with clinical factors

The preS deletion frequency and its association with each clinical factors is shown in Fig 2. Based on the FIB-4 values, the patients were stratified into three groups (over 3.45; n = 30,



**Fig 1. HBV preS deletion mapping using deep sequencing.** Deep sequencing of the region of the HBV genome encoding the 174 preS amino acids was carried out and the deleted amino acids were mapped in all 90 patients. Upper panel, the frequency of deletion of each amino acid position in the preS region is expressed for each patient as a color gradient between green and red. In this figure, these sequences are arranged in the order of the FIB-4 values. Lower panel, the average deletion frequency of the preS region per patient is shown.

<https://doi.org/10.1371/journal.pone.0212559.g001>

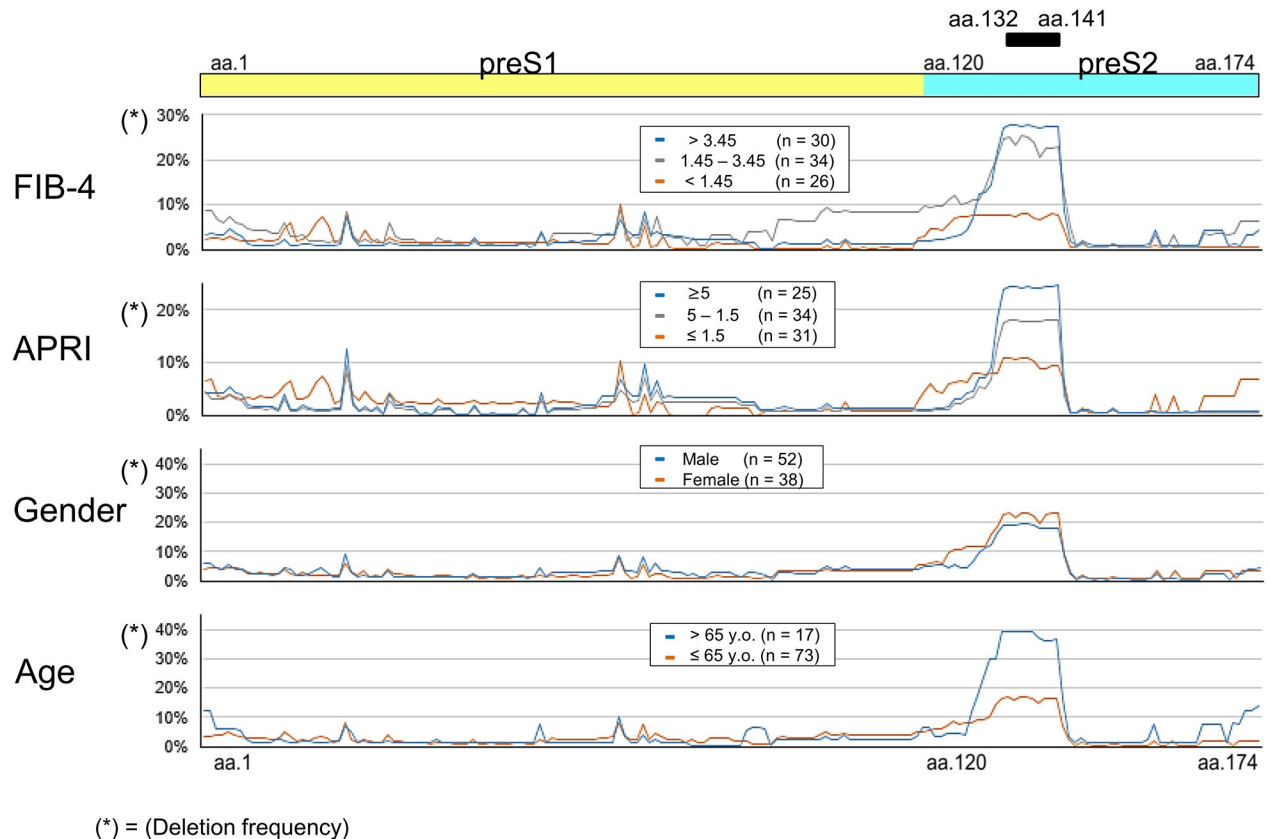
1.45–3.45;  $n = 34$ , and  $< 1.45$ ;  $n = 26$ ) and significant differences were observed in aa.132 to aa.141. Specifically, the preS region encoding aa.132 to aa.141 was deleted frequently in patients with a high FIB-4 index. The association of preS deletion frequency with APRI, gender and age was investigated (Fig 2), and it was also found that the preS region encoding aa.132 to aa.141 was deleted frequently in patients with a high aspartate aminotransferase—to—platelet ratio (APRI, and indicator of liver fibrosis) index and those who were elderly, while no association was found with gender.

### Deletion frequency of the preS region and its association with viral markers

The preS deletion frequency and its association with viral markers is shown in Fig 3. The preS region encoding aa.132 to aa.141 was deleted frequently in patients with high serum HBV DNA levels, high serum HBcrAg, preS1 start codon mutations and preS2 start codon mutations. Regarding the association with HBsAg, the preS region encoding aa.132 to aa.141 was deleted frequently in those with low HBsAg titers (Fig 3).

### Stratification of patients according to the frequency of preS codon 132–141 deletions

Because the data above showed a significant impact of the preS2 codon 132–141 region on clinical, as well as viral markers, we stratified the patients into three groups, according to the



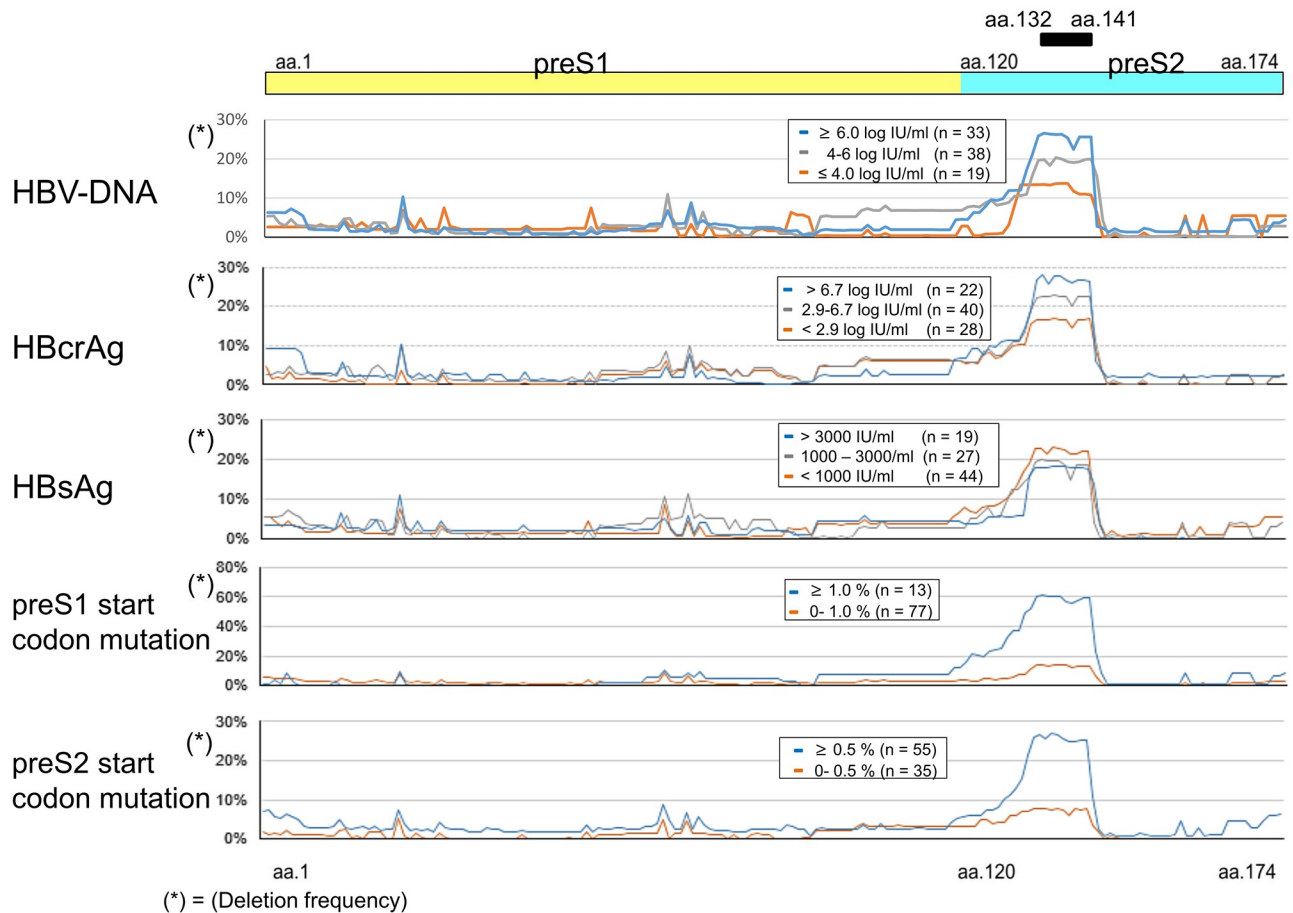
**Fig 2. PreS deletion frequency and its association with clinical factors.** PreS deletion frequency and its association with gender, age and clinical factors (FIB-4 and APRI). A. The patients were stratified into three groups according to their FIB-4 values (over 3.45;  $n = 30$ , 1.45–3.45;  $n = 34$ , and  $< 1.45$ ;  $n = 26$ ). B. The patients were stratified into three groups ( $5 \leq$ ;  $n = 25$ ,  $5 - 1.5$ ;  $n = 34$ , and  $\leq 1.5$ ;  $n = 31$ ) according to their APRI values. C. The patients were divided into two groups according to their sex (male;  $n = 52$ , female;  $n = 38$ ). D. The patients were stratified into two groups according to their age ( $> 65$  y.o.,  $n = 17$ ,  $\leq 65$  y.o.,  $n = 73$ ).

<https://doi.org/10.1371/journal.pone.0212559.g002>

frequency of codon 132-141 deletions. Group 0–0.1% includes 27 patients with 0% to  $< 0.1\%$  deletions; these patients are considered not to have preS codon 132-141 deletions because the deletion rate was below the deep sequencing cut-off. Group 0.1–10% includes 34 patients with  $\geq 0.1\%$  to  $< 10\%$  deletions and Group 10–100% includes 29 patients with  $\geq 10\%$  to 100% deletions. Thus, 63/90 (70%) of the patients had preS codon 132-141 deletions to some degree.

In Fig 4, the correlations between these three groups and clinical factors (age, AFP, and FIB-4) and viral markers (HBV DNA, HBcrAg, genotype, preS1/S2 start codon mutations) were investigated. As shown here, significant and stepwise associations with the preS codon 132-141 deletion were observed (Fig 4). It is noteworthy that a very low percentage of preS1/preS2 start codon mutations ( $> 1\%$  in preS1 and  $> 0.5\%$  in preS2) was significantly correlated with the preS codon 132-141 deletions; this would not have been evident without deep sequencing analysis. Genotypes were also significantly step-wisely associated with preS codon 132-141 deletions. While all patients (100%, 29/29) in 10–100% were infected with genotype C, 67% (18/27) in group 0–0.1% patients and 85% (29/34) in group 0.1–10% patients were infected with genotype C.

In S1 Fig a and b, preS deletion map and viral markers in each patient are demonstrated as lists in group 10–100% (a,  $n = 29$ ) and in group 0–0.1% (b,  $n = 27$ ).



**Fig 3. PreS deletion frequency and its association with viral markers.** The preS deletion frequency and its association with viral markers (HBV DNA, HBcrAg, HBsAg, preS1 start codon mutation, and preS2 start codon mutation). A, The patients were stratified into three groups according to their HBV DNA titers ( $\geq 6.0$  log IU/ml,  $n = 33$ , 4.0–6.0 log IU/ml,  $n = 38$  and  $\leq 4.0$  log IU/ml,  $n = 19$ ). B, The patients were stratified into three groups according to their HBcrAg titers ( $> 6.7$  log IU/ml,  $n = 22$ , 2.9–6.7 log IU/ml,  $n = 40$  and  $< 2.9$  log IU/ml,  $n = 28$ ). C, The patients were stratified into three groups according to their HBsAg titers ( $> 3000$  IU/ml,  $n = 19$ , 1000–3000 IU/ml,  $n = 27$  and  $< 1000$  IU/ml,  $n = 44$ ). D, The patients were stratified into two groups according to the frequency of preS1 start codon mutations ( $\geq 1.0\%$ ,  $n = 13$  and 0.0–1.0%,  $n = 77$ ). E, The patients were stratified into two groups according to the frequency of preS2 start codon mutations ( $\geq 0.5\%$ ,  $n = 55$  and 0.0–0.5%,  $n = 35$ ).

<https://doi.org/10.1371/journal.pone.0212559.g003>

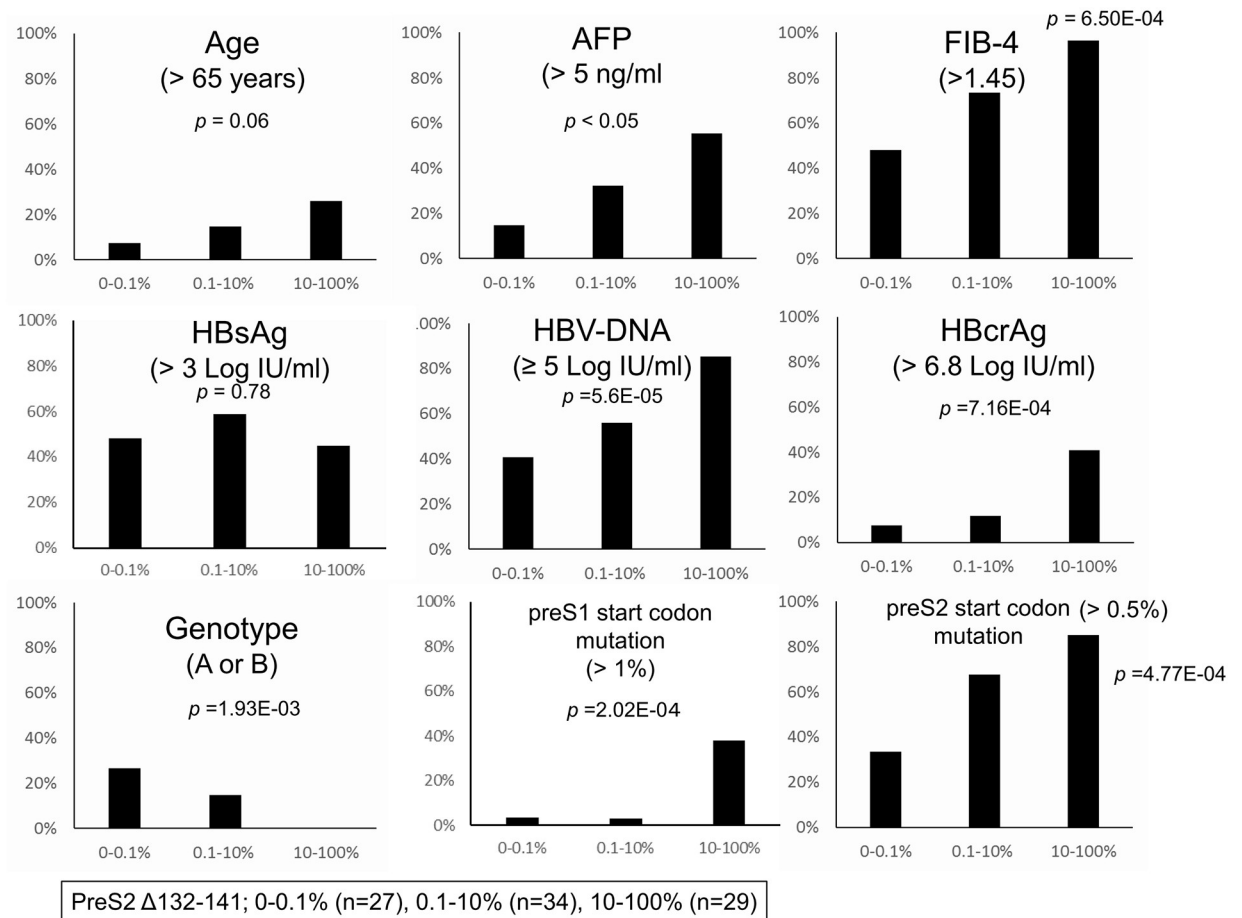
### Multivariate analysis of independent factors associated with preS codon 132–141 deletions

It was clear that the preS codon 132–141 deletion rate was significantly associated with various factors, therefore, we carried out multivariate analysis to identify factors influenced independently by the deletions (Table 2).

Multivariate analysis extracted FIB-4, HBcrAg and preS1 start codon mutation as independent factors.

### Impact of preS codon 132–141 deletions on the serum HBsAg and HBcrAg titer, relative to the serum HBV DNA level

We investigated the influence of preS codon 132–141 deletions on the serum HBV markers, focusing on HBsAg and HBcrAg. Because our interest was to determine how production of these viral antigens from the viral DNA is affected by the presence of preS codon 132–141 deletion, they are expressed as the amount of viral antigen relative to the serum HBV DNA titer:



**Fig 4. The correlations between the frequency of preS codon 132–141 deletions and host factors and viral markers.** The correlations between the frequency of preS codon 132-141 deletions and host factors (age, AFP, and FIB-4) and viral markers (HBV DNA, HBcrAg, genotype and preS1/S2 start codon mutations) are demonstrated.

<https://doi.org/10.1371/journal.pone.0212559.g004>

HBsAg/ HBV DNA and HBcrAg/ HBV DNA. As shown in Fig 5, left panel, the HBsAg titer relative to the HBV DNA value was significantly lower in cases with high frequencies (10–100%) of deletions than those with low frequencies. Conversely, the HBcrAg titer relative to the HBV DNA value was significantly higher in cases with high frequencies (10–100%) of deletions than those with low frequencies (0–0.1%, Fig 5, right panel).

### Future NA therapy in patients with preS codon 132–141 deletions

In order to determine the influence of the preS codon 132–141 deletions on the requirement for future NA therapy, Kaplan-Meier curves were drawn according to the frequency of the preS codon 132–141 deletions. As shown in Fig 6, compared to the patients with the preS codon 132–141 deletions 0–0.1%, NA therapy was introduced more often in patients with high deletion frequencies (10–100%,  $p = 0.0095$ ) or with intermediate deletion frequencies (1–10%,  $p = 0.018$ ), over the ensuing five years (Fig 6).

### Discussion

In this study, based on deep sequencing analysis, we have shown that deletions in the preS region of the HBV genome are present in the viral quasispecies of HBV-infected Japanese

**Table 2. Univariate and multivariate analysis for factors associated with preS codon 132–141 deletion ( $\geq 10\%$ ).**

	Univariate analysis				p-value	Multivariate analysis		
	0–10% (n = 61)		10–100% (n = 29)			OR	95% CI	p-value
	n	(%)	n	(%)				
Age (> 65 years)	8	(13%)	9	(31%)	< 0.05*	2.1	(0.5–8.6)	0.32
Gender (Male)	37	(61%)	15	(52%)	0.28			
Alb ( $\leq 4.0$ g/dL)	20	(33%)	10	(34%)	0.66			
ALT ( $\geq 31$ IU/mL)	31	(51%)	16	(55%)	0.44			
AFP ( $\geq 5$ ng/ml)	16	(26%)	15	(52%)	< 0.05*	1.0	(0.3–4.0)	0.95
FIB-4 (> 1.45)	38	(62%)	26	(90%)	< 0.01**	9.0	(1.2–65.6)	< 0.05*
HCC (yes)	10	(16%)	5	(17%)	0.66			
HBsAg (> 3.0 LogIU/mL)	33	(54%)	13	(45%)	0.50			
HBV-DNA ( $\geq 5.0$ LogIU/ml)	30	(49%)	23	(79%)	< 0.01**	0.8	(0.2–3.3)	0.71
HBcrAg (> 6.8 LogIU/ml)	6	(10%)	11	(38%)	< 0.01**	6.8	(1.2–39.5)	< 0.05*
HBeAg (positive)	9	(15%)	14	(48%)	< 0.01**	1.4	(0.3–7.6)	0.66
Genotype (non-C)‡	14	(23%)	0	(0%)	< 0.01**			
preS1 start codon mutation (> 1%)	2	(3%)	11	(38%)	< 0.01**	28.6	(3.5–236.8)	< 0.01**
preS2 start codon mutation (> 0.5%)	32	(52%)	23	(79%)	< 0.01**	1.5	(0.9–2.3)	0.10

Univariate analysis and multivariate were performed with Fisher exact t-test and logistic -regression analysis respectively.

‡Genotype was excluded from multivariate analysis because calculation was impossible due to deviated distribution of patients.

\* < 0.05,

\*\* < 0.01

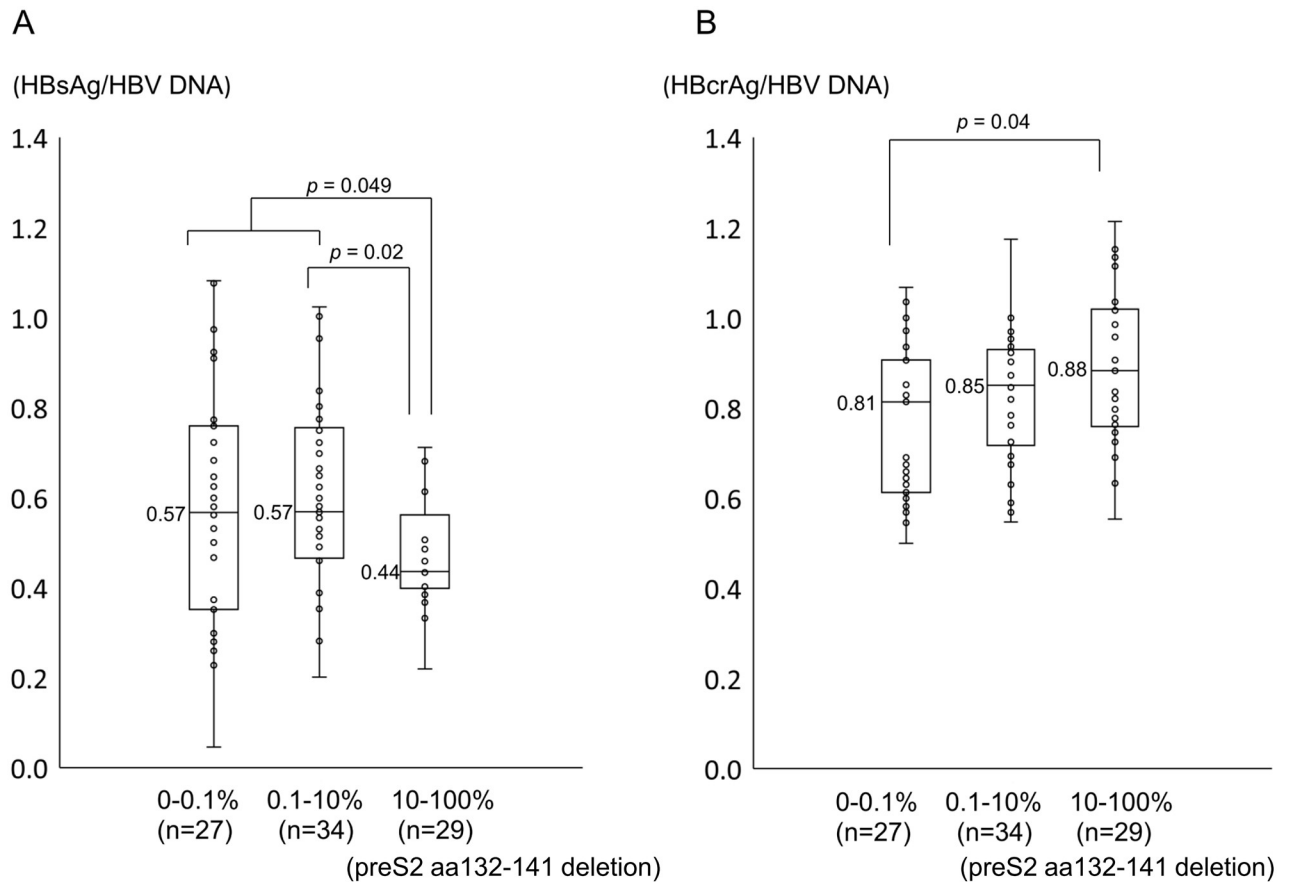
<https://doi.org/10.1371/journal.pone.0212559.t002>

patients and that deletions of the region encoding codon 132–141 are most frequent. Importantly, the presence of these deletions showed significant positive correlations with fibrosis (FIB-4 and APRI) and high AFP values. Moreover, the presence of these deletions was correlated with viral markers (the HBV DNA titer and HBcrAg titer) and preS1/S2 start codon mutations, even when present as a minor species. From multivariate analysis, advanced fibrosis (FIB-4), high HBcrAg titer, and a high prevalence of preS1 start codon mutations were independently correlated with preS2 codon 132–141 deletions.

It has been reported that preS mutants, including deletion mutants, are found frequently in patients with advanced liver disease, including liver cirrhosis and HCC [6, 7, 15–19] and the mutants have even been found to be correlated with NA resistance [20]. These mutations are considered to be selected as the result of immune pressure, because many T/B-cell epitopes are present in the preS region [6]. On the other hand, although various preS deletion mutants have been reported [6, 7, 10, 21], the most relevant locus in the region affecting clinical characteristics, and the frequency its deletion in individual patients, has been unclear because most of those previous studies, using direct sequencing or sequencing of clones, could not quantify the mutation frequencies. In the real world of chronic viral infection, viruses are considered to infect as a population of different but closely related genomes, the so-called viral quasispecies, and, therefore, preS mutants may be present at various frequencies [22]. In this study, using deep sequencing, we showed that the preS regions also formed a quasispecies in a single patient. By investigating the deletion frequencies, deletion of the region encoding preS codon 132–141 was shown to be the most frequent and common among the patients.

What is the role of preS codon 132–141 deletions in the pathogenesis of HBV-induced liver disease? In previous studies, preS deletions around codon 132–141 were found frequently in patients with liver biopsy specimens with type II ground-glass hepatocytes (GGH) [4, 9, 19, 23]. Because the preS region around codon 132–141 overlaps a CTL epitope, such deletions



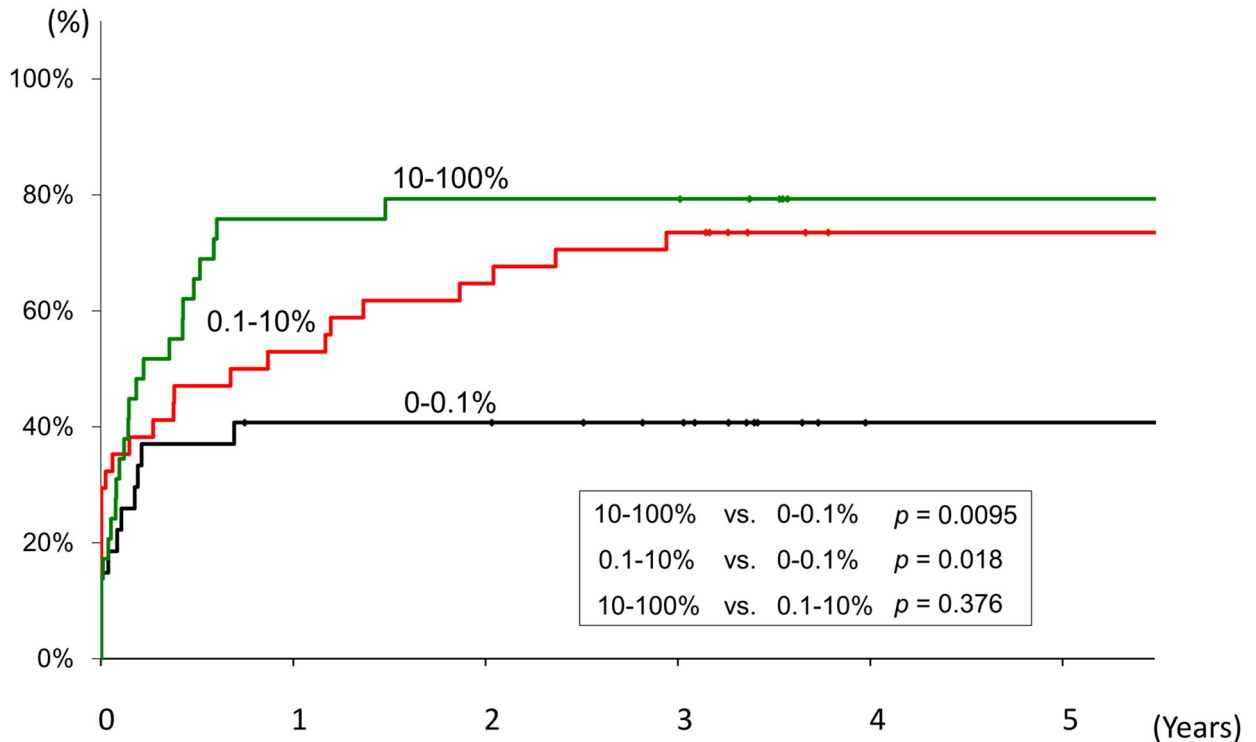


**Fig 5. The impact of preS codon 132–141 deletions on the serum HBsAg and HBcrAg titers, relative to the serum HBV DNA level.** Left panel, the HBsAg titer relative to the HBV DNA value was significantly lower in cases with high frequencies of mutations (10–100%) than in those with low frequencies. Right panel, the HBcrAg titer relative to the HBV DNA value was significantly higher in cases with high frequencies of mutations (10–100%) compared to those with low frequencies (0–0.1%).

<https://doi.org/10.1371/journal.pone.0212559.g005>

might occur as a result of immune escape and, therefore, it is plausible that fibrosis is independently associated with codon 132–141 deletions, because advanced fibrosis often is the results of significant CTL-induced inflammation. On the other hand, these deletions may result in the accumulation of HBsAg in the endoplasmic reticulum (ER), when HBsAg secretion is diminished [10, 23]. In such a state, liver damage might be associated with ER stress [24, 25], although cellular damage unrelated to ER stress could also occur [4, 5, 19, 26–28]. Because type-II GGH is often found in advanced liver disease, and is even considered as a precancerous lesion, the results of our study are compatible with those previous studies and rather enhance the correlation of mutations in the viral genome and disease progression.

In this study, it was also demonstrated that the HBV-DNA and HBcrAg titers paralleled the frequency of preS codon 132–141 deletions. Although the mechanism of the correlation remains unclear, it is also plausible that highly replication-competent viruses with high cccDNA transcription activity develop immune-escape mutations more rapidly to escape from hosts' immune attack. Though high HBV-DNA and HBcrAg titers both reflect high intrahepatic HBV cccDNA titers, as described earlier, HBcrAg might reflect intrahepatic HBV-cccDNA transcription activity more accurately, considering that multivariate analysis showed HBcrAg to be independently associated with the deletions. A recent study also reported the significant



**Fig 6. Frequency of the preS codon 132–141 deletions and its influence on the requirement for future nucleoside analogue therapy.** Kaplan-Meier curves were drawn according to the frequency of the preS codon 132–141 deletions to determine the influence of the deletions on the requirement for future nucleoside analogue therapy over the ensuing five years. As is shown, compared to the patients with the preS codon 132–141 deletions 0–0.1%, NA therapy was introduced more often in patients with high deletion frequencies (10–100%,  $p = 0.0095$ ) or with intermediate deletion frequencies (1–10%,  $p = 0.018$ ).

<https://doi.org/10.1371/journal.pone.0212559.g006>

association between HBcrAg and intrahepatic HBV-cccDNA transcription activity [29]. On the other hand, HBsAg, another viral marker which also reflects intrahepatic HBV cccDNA [11, 30, 31], was not extracted in this study as a factor associated with the deletions. Though its reasons are not clear, if high HBcrAg titer (or high HBV-DNA titer) reflects robust intrahepatic HBV-cccDNA transcription activity as recently reported [29], HBsAg production per HBV-cccDNA might be even decreased when compared with the production of HBcrAg or HBV-DNA in the presence of preS deletions. In this sense, we re-evaluated HBsAg and HBcrAg titer after correction by HBV-DNA titer (Fig 5). The result of Fig 5a might reflect association among preS deletions, intrahepatic HBsAg accumulation and diminished extracellular HBsAg secretion observed in in-vitro studies [10] while Fig 5b might reflect robust intrahepatic HBV-cccDNA transcription activity of HBcrAg though further studies are needed.

Interestingly, in our study, preS1 start codon mutation was also extracted as an independent variable in its association with the preS2 codon 132–141 deletions, even when present as minor species (> 1%). Although the reason for this is unknown, conservation of the preS1 start codon, which is indispensable for large S antigen production, might not be needed because the large S protein lacking 10 amino acids (codon 132–141) produced in the presence of the preS2 codon 132–141 might have dysfunction. Likewise, conservation of the preS2 start codon, which is indispensable for middle S antigen production might not be needed though its correlation with the deletion might be rather weak. Therefore, we speculate that preS1 (and preS2) start codon mutation might not be a cause but rather a result of the preS2 codon 132–141 deletion, although further studies are needed.

One of the limitations of the study is that we could not obtain liver tissue from all of the patients and, therefore, could not investigate the correlation of serum HBsAg, HBcrAg, HBV DNA and the preS codon 132–141 deletions with intrahepatic HBV and liver histology. Because the ultimate goal of anti-HBV therapy is the elimination of HBV from the liver, the association of those markers with liver histology, HBV cccDNA titer and the integration of viral DNA into chromosomal DNA would reveal further the status of HBV-induced liver disease.

In conclusion, quasispecies analysis of preS deletions in the HBV genome using deep sequencing revealed a close correlation between these deletions and the state of liver disease, as well as HBV markers, even when they are present as minor populations. Understanding the interrelationship among those viral markers in association with the state of liver disease would further advance our understanding of the mechanisms of liver disease progression.

## Patients and methods

### Patients

Ninety patients chronically infected with HBV and followed-up at Yamanashi University Hospital after 2004 were included in the study. In order to include patients with advanced diseases as well as inactive carriers, 61 consecutive patients clinically diagnosed as cirrhosis or chronic hepatitis and 29 consecutive patients clinically diagnosed as inactive carriers were enrolled. All patients also met the following criteria: (1) Had not received NA therapy previously at the time of enrollment. (2) Serum was available for HBV sequence analysis. (3) Were hepatitis C antibody negative. (4) Had no other forms of hepatitis, such as primary biliary cirrhosis, autoimmune liver disease or alcoholic liver disease. (5) Were free of co-infection with human immunodeficiency virus. All enrolled patients were positive for HBsAg. HBV DNA was measured by the Quantiplex HBV DNA assay (Bayer Diagnostics, Emeryville, CA, USA), transcription-mediated amplification assay (Chugai Diagnostics Science Co., Ltd., Tokyo, Japan), or COBAS Amplicor HBV Monitor Test v2.0 (Roche Diagnostics, Indianapolis, IN, USA). The clinical characteristics of these patients are shown in [Table 1](#).

We firstly intended to include larger number of patients in this study. However, at that time, NA therapy was introduced in Japan, and NA treatment was started for most active hepatitis (chronic hepatitis and liver cirrhosis) patients. Therefore, it became difficult to include more patients, especially patients with active disease.

### Ethics statement

Informed consent was obtained for participation in the study protocol, which had been approved by the Human Ethics Review Committee of Yamanashi University. All included patients were adult. Written informed consent or ethics committee approved opt-out consent was obtained from all individual patients included in the study.

### DNA extraction, PCR, and deep sequencing

The preS region of HBV DNA was amplified from the patients' sera using two-step PCR. The first and the second-round primers are shown in the [S1 Table](#). Briefly, viral DNA was extracted from stored sera using QIAamp MinElute Virus Spin Kits (QIAGEN, Tokyo, Japan) with QIAcube (QIAGEN), and then the extracted DNA was subjected to two-step PCR. The primers for the second-round PCR had bar codes 10 nucleotides in length attached and these differed for each sample, so that the PCR products from each sample were identifiable ([S1 Table](#)). After the PCR products were quantified using a Pico Green dsDNA Assay Kit (Invitrogen, Tokyo,

Japan), the concentrations of the samples were adjusted to a common value and the samples were pooled.

Libraries were then subjected to emulsion PCR, the enriched DNA beads were loaded onto a picotiter plate and pyrosequencing was carried out with a Roche GS Junior/454 sequencing system using titanium chemistry (Roche, Branford, CT). The Roche Variant Analyzer version 2.5pl (Roche) and Microsoft Excel (Microsoft, Tokyo, Japan) were used for the analysis. The method for pyrosequencing was described previously in more detail [32].

### Detection of mutations and deletions in preS1 and preS2 regions using deep sequencing

Deep sequencing involved pyro-sequencing of the 522 nt PCR amplified preS1/S2 region. The average depth was approximately 2279 reads and the proportion of preS1 and preS2 start codon mutations and preS1 and preS2 deletions was determined for each patient. The cut-off for the presence of mutations was set at 0.1%, based on a control experiment using the plasmid template.

### Statistical analysis

Statistical differences in the parameters, including all available demographic, biochemical, hematological, and virological variables, were determined for the various patient groups by Fisher's exact probability test for categorical variables. The odds ratios and 95% confidence intervals were calculated. The Cochran-Armitage trend test was used to look for trends in the categorical data among three groups, divided on the basis of differences in the preS deletion frequency. A logistic regression analysis was used to investigate independent variables associated with the preS codon 132–141 deletion. In order to evaluate the requirement for future NA treatment, Kaplan-Meier curves were drawn and the log-rank test was performed. P values of <0.05 by the two-tailed test were considered to indicate statistical significance.

### Supporting information

**S1 Table. The first and the second-round primers for the preS hepatitis B virus genomic region.** The primers for the second-round PCR had bar codes 10 nucleotides in length attached and these differed for each sample, so that the PCR products from each sample were identifiable.

(XLSX)

**S1 Fig. preS deletion map and viral markers in group 10–100% and group 0–0.1%.** In [S1 Fig](#), preS deletion map and viral markers in each patient are demonstrated as lists in group 10–100% (1a, n = 29) and in group 0–0.1% (1b, n = 27).

(PDF)

### Author Contributions

**Conceptualization:** Yuichiro Suzuki, Shinya Maekawa, Nobutoshi Komatsu.

**Data curation:** Yuichiro Suzuki, Shinya Maekawa.

**Formal analysis:** Shinya Maekawa.

**Investigation:** Yuichiro Suzuki, Shinya Maekawa, Nobutoshi Komatsu, Mitsuaki Sato, Akihisa Tatsumi, Mika Miura, Shuya Matsuda, Masaru Muraoka, Shinichi Takano, Minoru Sakamoto.

**Methodology:** Yuichiro Suzuki, Shinya Maekawa, Tatsuya Yamaguchi, Taisuke Inoue, Tadashi Sato, Minoru Sakamoto.

**Supervision:** Atsuya Yamashita, Kohji Moriishi.

**Writing – original draft:** Yuichiro Suzuki, Shinya Maekawa.

**Writing – review & editing:** Natsuko Nakakuki, Fumitake Amemiya, Mitsuharu Fukasawa, Yasuhiro Nakayama, Nobuyuki Enomoto.

## References

1. Organization WH. Global hepatitis report 2017: executive summary. 2017.
2. Pollicino T, Cacciola I, Saffioti F, Raimondo G. Hepatitis B virus PreS/S gene variants: pathobiology and clinical implications. *J Hepatol*. 2014; 61(2):408–17. Epub 2014/05/08. <https://doi.org/10.1016/j.jhep.2014.04.041> PMID: 24801416.
3. Sugauchi F, Ohno T, Orito E, Sakugawa H, Ichida T, Komatsu M, et al. Influence of hepatitis B virus genotypes on the development of preS deletions and advanced liver disease. *J Med Virol*. 2003; 70(4):537–44. Epub 2003/06/10. <https://doi.org/10.1002/jmv.10428> PMID: 12794715.
4. Hsieh YH, Su IJ, Wang HC, Chang WW, Lei HY, Lai MD, et al. Pre-S mutant surface antigens in chronic hepatitis B virus infection induce oxidative stress and DNA damage. *Carcinogenesis*. 2004; 25(10):2023–32. Epub 2004/06/08. <https://doi.org/10.1093/carcin/bgh207> PMID: 15180947.
5. Wang HC, Chang WT, Chang WW, Wu HC, Huang W, Lei HY, et al. Hepatitis B virus pre-S2 mutant upregulates cyclin A expression and induces nodular proliferation of hepatocytes. *Hepatology*. 2005; 41(4):761–70. Epub 2005/02/24. <https://doi.org/10.1002/hep.20615> PMID: 15726643.
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(4):1153–68. Epub 2006/04/19. <https://doi.org/10.1053/j.gastro.2006.01.011> PMID: 16618410.
7. Chen CH, Hung CH, Lee CM, Hu TH, Wang JH, Wang JC, et al. Pre-S deletion and complex mutations of hepatitis B virus related to advanced liver disease in HBeAg-negative patients. *Gastroenterology*. 2007; 133(5):1466–74. Epub 2007/10/05. <https://doi.org/10.1053/j.gastro.2007.09.002> PMID: 17915220.
8. Gao ZY, Li T, Wang J, Du JM, Li YJ, Li J, et al. Mutations in preS genes of genotype C hepatitis B virus in patients with chronic hepatitis B and hepatocellular carcinoma. *J Gastroenterol*. 2007; 42(9):761–8. Epub 2007/09/19. <https://doi.org/10.1007/s00535-007-2085-1> PMID: 17876546.
9. 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(8 Pt 1):1169–74. Epub 2008/05/29. <https://doi.org/10.1111/j.1440-1746.2008.05348.x> PMID: 18505413.
10. Pollicino T, Amadio G, Restuccia A, Raffa G, Alibrandi A, Cutroneo G, et al. Impact of hepatitis B virus (HBV) preS/S genomic variability on HBV surface antigen and HBV DNA serum levels. *Hepatology*. 2012; 56(2):434–43. Epub 2012/01/25. <https://doi.org/10.1002/hep.25592> PMID: 22271491.
11. Chan HL, Thompson A, Martinot-Peignoux M, Piratvisuth T, Cornberg M, Brunetto MR, et al. Hepatitis B surface antigen quantification: why and how to use it in 2011—a core group report. *J Hepatol*. 2011; 55(5):1121–31. Epub 2011/07/02. <https://doi.org/10.1016/j.jhep.2011.06.006> PMID: 21718667.
12. Rokuhara A, Tanaka E, Matsumoto A, Kimura T, Yamaura T, Orii K, et al. Clinical evaluation of a new enzyme immunoassay for hepatitis B virus core-related antigen; a marker distinct from viral DNA for monitoring lamivudine treatment. *J Viral Hepat*. 2003; 10(4):324–30. Epub 2003/06/26. PMID: 12823601.
13. Maasoumy B, Wiegand SB, Jaroszewicz J, Bremer B, Lehmann P, Deterding K, et al. Hepatitis B core-related antigen (HBcrAg) levels in the natural history of hepatitis B virus infection in a large European cohort predominantly infected with genotypes A and D. *Clin Microbiol Infect*. 2015; 21(6):606 e1–10. Epub 2015/02/24. <https://doi.org/10.1016/j.cmi.2015.02.010> PMID: 25700889.
14. Suzuki F, Miyakoshi H, Kobayashi M, Kumada H. Correlation between serum hepatitis B virus core-related antigen and intrahepatic covalently closed circular DNA in chronic hepatitis B patients. *J Med Virol*. 2009; 81(1):27–33. Epub 2008/11/26. <https://doi.org/10.1002/jmv.21339> PMID: 19031469.
15. Bai X, Jia J-a, Fang M, Chen S, Liang X, Zhu S, et al. Deep sequencing of HBV pre-S region reveals high heterogeneity of HBV genotypes and associations of word pattern frequencies with HCC. *PLoS genetics*. 2018; 14(2):e1007206. <https://doi.org/10.1371/journal.pgen.1007206> PMID: 29474353
16. Fang ZL, Sabin CA, Dong BQ, Wei SC, Chen QY, Fang KX, et al. Hepatitis B virus pre-S deletion mutations are a risk factor for hepatocellular carcinoma: a matched nested case-control study. *J Gen Virol*.

- 2008; 89(Pt 11):2882–90. Epub 2008/10/22. <https://doi.org/10.1099/vir.0.2008/002824-0> PMID: 18931087
17. Li YW, Yang FC, Lu HQ, Zhang JS. Hepatocellular carcinoma and hepatitis B surface protein. *World J Gastroenterol.* 2016; 22(6):1943–52. Epub 2016/02/16. <https://doi.org/10.3748/wjg.v22.i6.1943> PMID: 26877602
  18. Liu S, Zhang H, Gu C, Yin J, He Y, Xie J, et al. Associations between hepatitis B virus mutations and the risk of hepatocellular carcinoma: a meta-analysis. *J Natl Cancer Inst.* 2009; 101(15):1066–82. Epub 2009/07/04. <https://doi.org/10.1093/jnci/djp180> PMID: 19574418
  19. Su IJ, Wang LH, Hsieh WC, Wu HC, Teng CF, Tsai HW, et al. The emerging role of hepatitis B virus pre-S2 deletion mutant proteins in HBV tumorigenesis. *J Biomed Sci.* 2014; 21:98. Epub 2014/10/16. <https://doi.org/10.1186/s12929-014-0098-7> PMID: 25316153
  20. Sueki R, Maekawa S, Miura M, Kadokura M, Komase K, Shindo H, et al. Correlation between pretreatment viral sequences and the emergence of lamivudine resistance in hepatitis B virus infection. *J Med Virol.* 2012; 84(9):1360–8. Epub 2012/07/25. <https://doi.org/10.1002/jmv.23314> PMID: 22825814.
  21. Zhang D, Dong P, Zhang K, Deng L, Bach C, Chen W, et al. Whole genome HBV deletion profiles and the accumulation of preS deletion mutant during antiviral treatment. *BMC Microbiol.* 2012; 12:307. Epub 2013/01/01. <https://doi.org/10.1186/1471-2180-12-307> PMID: 23272650
  22. Rodriguez-Frias F, Buti M, Tabernero D, Homs M. Quasispecies structure, cornerstone of hepatitis B virus infection: mass sequencing approach. *World J Gastroenterol.* 2013; 19(41):6995–7023. Epub 2013/11/14. <https://doi.org/10.3748/wjg.v19.i41.6995> PMID: 24222943
  23. Wang HC, Wu HC, Chen CF, Fausto N, Lei HY, Su IJ. Different types of ground glass hepatocytes in chronic hepatitis B virus infection contain specific pre-S mutants that may induce endoplasmic reticulum stress. *Am J Pathol.* 2003; 163(6):2441–9. Epub 2003/11/25. [https://doi.org/10.1016/S0002-9440\(10\)63599-7](https://doi.org/10.1016/S0002-9440(10)63599-7) PMID: 14633616
  24. Chua PK, Wang RY, Lin MH, Masuda T, Suk FM, Shih C. Reduced secretion of virions and hepatitis B virus (HBV) surface antigen of a naturally occurring HBV variant correlates with the accumulation of the small S envelope protein in the endoplasmic reticulum and Golgi apparatus. *J Virol.* 2005; 79(21):13483–96. Epub 2005/10/18. <https://doi.org/10.1128/JVI.79.21.13483-13496.2005> PMID: 16227269
  25. Wang HC, Huang W, Lai MD, Su IJ. Hepatitis B virus pre-S mutants, endoplasmic reticulum stress and hepatocarcinogenesis. *Cancer Sci.* 2006; 97(8):683–8. Epub 2006/07/26. <https://doi.org/10.1111/j.1349-7006.2006.00235.x> PMID: 16863502.
  26. Hsieh YH, Chang YY, Su IJ, Yen CJ, Liu YR, Liu RJ, et al. Hepatitis B virus pre-S2 mutant large surface protein inhibits DNA double-strand break repair and leads to genome instability in hepatocarcinogenesis. *J Pathol.* 2015; 236(3):337–47. Epub 2015/03/18. <https://doi.org/10.1002/path.4531> PMID: 25775999.
  27. 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(2):466–72. Epub 2011/12/14. <https://doi.org/10.1093/carcin/bgr296> PMID: 22159224.
  28. Zhang X, Gao L, Liang X, Guo M, Wang R, Pan Y, et al. HBV preS2 transactivates FOXP3 expression in malignant hepatocytes. *Liver Int.* 2015; 35(3):1087–94. Epub 2014/07/23. <https://doi.org/10.1111/liv.12642> PMID: 25047684.
  29. Testoni B, Lebossé F, Scholtes C, Berby F, Miaglia C, Subic M, et al. Serum Hepatitis B core-related antigen (HBcrAg) correlates with covalently-closed circular DNA transcriptional activity in chronic hepatitis B patients. *J Hepatol.* 2018.
  30. Chan HL, Wong VW, Tse AM, Tse CH, Chim AM, Chan HY, et al. Serum hepatitis B surface antigen quantitation can reflect hepatitis B virus in the liver and predict treatment response. *Clin Gastroenterol Hepatol.* 2007; 5(12):1462–8. Epub 2007/12/07. <https://doi.org/10.1016/j.cgh.2007.09.005> PMID: 18054753.
  31. Lin LY, Wong VW, Zhou HJ, Chan HY, Gui HL, Guo SM, et al. Relationship between serum hepatitis B virus DNA and surface antigen with covalently closed circular DNA in HBsAg-negative patients. *J Med Virol.* 2010; 82(9):1494–500. Epub 2010/07/22. <https://doi.org/10.1002/jmv.21863> PMID: 20648602.
  32. Sato M, Maekawa S, Komatsu N, Tatsumi A, Miura M, Muraoka M, et al. Deep sequencing and phylogenetic analysis of variants resistant to interferon-based protease inhibitor therapy in chronic hepatitis induced by genotype 1b hepatitis C virus. *J Virol.* 2015; 89(11):6105–16. Epub 2015/03/27. <https://doi.org/10.1128/JVI.03127-14> PMID: 25810555