RESEARCH ARTICLE

Open Access

CrossMark

Quasispecies variant of pre-S/S gene in HBV-related hepatocellular carcinoma with HBs antigen positive and occult infection

Yuri Hatazawa[1], Yoshihiko Yano[1,2*], Rina Okada[1], Toshihito Tanahashi[3], Hiroki Hayashi[1], Hirotaka Hirano[1], Akihiro Minami[1], Yuki Kawano[1], Motofumi Tanaka[4], Takumi Fukumoto[4], Yoshiki Murakami[5], Masaru Yoshida[1] and Yoshitake Hayashi[2]

Abstract

Background: Hepatocellular carcinoma (HCC) can develop in patients who are negative for the hepatitis B surface antigen (HBsAg) in serum but positive for hepatitis B virus (HBV) DNA in the liver, referred to as occult HBV infection (OBI). Previous reports showed that HBV variants in OBI-related HCC are different from those in HBsAg-positive HCC. In the present study, HBV quasispecies based on the pre-S/S gene in OBI-related HCC patients were examined by high throughput sequencing and compared with those in HBsAg-positive HCC.

Methods: Nineteen tissue samples (9 OBI-related and 10 HBsAg-positive non-cancerous tissues) were collected at the time of surgery at Kobe University Hospital. The quasispecies with more than 1% variation in the pre-S/S region were isolated and analysed by ultra-deep sequencing.

Results: There were no significant differences in the major HBV populations, which exhibit more than 20% variation within the entire pre-S/S region, between OBI-related HCC and HBsAg-positive HCC. However, the prevalences of major populations with pre-S2 region mutations and of minor populations with polymerized human serum albumin-binding domain mutations were significantly higher in OBI-related HCC than in HBsAg-positive HCC. Moreover, the major variant populations associated with the B-cell epitope, located within the pre-S1 region, and the a determinant domain, located in the S region, were detected frequently in HBsAg-positive HCC. The minor populations of variants harbouring the W4R, L30S, Q118R/Stop, N123D and S124F/P mutations in the pre-S region and the L21F/S and L42F/S mutations in the S region were detected more frequently in OBI-related HCC than in HBsAg-positive HCC.

Conclusions: Ultra-deep sequencing revealed that the B-cell epitope domain in the pre-S1 region and alpha determinant domain in the S region were variable in HBsAg-positive HCC, although the quasispecies associated with the pre-S2 region were highly prevalent in OBI-related HCC.

Trial registration: Ref: R000034382/UMIN000030113; Retrospectively registered 25 November 2017.

Keywords: HBV, Pre-S/S, Quasispecies, Occult, HCC, Ultra-deep sequencing

* Correspondence: yanoyo@med.kobe-u.ac.jp
[1]Division of Gastroenterology, Department of Internal Medicine, Kobe University Graduate School of Medicine, Kobe, Hyogo 650-0017, Japan
[2]Division of Molecular Medicine & Medical Genetics, Department of Pathology, Kobe University Graduate School of Medicine, Kobe, Japan
Full list of author information is available at the end of the article

Hatazawa *et al. Infectious Agents and Cancer* (2018) 13:7

Page 2 of 8

Background

It is estimated that approximately 248 million people worldwide are infected with the hepatitis B virus (HBV), with the highest prevalence occurring in Asia and Africa [1]. The clinical courses of chronic hepatitis B vary. Whereas some patients may progress from chronic hepatitis to cirrhosis and hepatocellular carcinoma (HCC) after HBV infection, others may naturally eliminate HBV from the blood and become hepatitis B surface antigen (HBsAg) negative [2, 3]. HBsAg is used as a marker to screen for HBV infection, and the HBsAg titre generally decreases with long-term immune responses in the host. The annual seroclearance rate of HBsAg is 0.7%–1.2% [4–6]. Recently, the goal of chronic hepatitis B therapy has been to achieve seroclearance of HBsAg [7]. However, once HBV infects the host, it integrates into the host genome and establishes a latent infection in the liver for life [8].

A serologically negative HBsAg but positive HBV DNA status in serum or the liver is referred to as occult HBV infection (OBI) [9, 10]. OBI is a potential risk factor for cirrhosis and HCC. Although it has been reported that HBV variants of the pre-S/S region are related to HCC generation and HBsAg seroclearance [11–13], the cause of carcinogenesis in OBI patients remains unclear.

The variation among viral quasispecies is conventionally analysed by direct sequencing. However, it is difficult to detect variation levels less than 20% using this method. The analysis of drug-resistance and disease-specific mutations in viruses has recently become possible following technical advances in next-generation sequencing [14, 15]. In particular, thousands of viral quasispecies within a single host can be analysed by ultra-deep sequencing methods, and it has been suggested that the detection of minor quasispecies populations are related to the clinical efficacy [16, 17].

This study focused on quasispecies based on the HBV pre-S/S gene in surgically resected HCC specimens. Using ultra-deep sequencing methods, the quasispecies were analysed and compared between OBI-related HCC and HBsAg-positive HCC. We investigated the specific mutations or variation patterns in OBI-related HCC tissues.

Methods

Patients and sample collection

Nineteen HCC patients who underwent surgical resection at Kobe University Hospital were enrolled in this study. Of the 19 patients, 9 (age 69.7 ± 5.6 years, 8 males and 1 female) were serologically negative for HBsAg, while 10 (age 53.0 ± 11.5 years, 9 males and 1 female) were positive for HBsAg. Laboratory data including platelet counts, transaminase levels, haemoglobin A1c,

hyaluronic acid and quantitative HBsAg levels (HBsAg-HQ: Fujirebio, Tokyo, Japan) at the time of surgery were retrieved from patients' medical records. In addition, hepatitis B core antibody (anti-HBc; Architect AUSAB, Abbott Japan, Tokyo, Japan) and HBV DNA levels (Cobas® TaqMan HBV Test, v2.0, Roche Diagnostics, Basel, Switzerland) were also examined. The cut-off levels of HBsAg-HQ and HBV DNA were estimated 5 mIU/ml and 2.06 log copies/ml, respectively. Patients negative for anti-HBc or positive for hepatitis C virus antibodies were excluded from this study.

HBV extraction and ultra-deep sequencing analysis

Viral DNA was extracted from 200 μg of non-cancer tissue using the QIAGEN DNA Mini Kit (Qiagen, Tokyo, Japan) according to the manufacturer's protocol. The pre-S/S region was amplified by nested PCR using specific primers [18]. The presence of an amplified product was confirmed by electrophoresis on 2% agarose gels, which were stained with ethidium bromide and visualized on a UV transilluminator. All of the necessary precautions were taken to prevent cross-contamination. A negative control was included in each assay.

Ultra-deep sequencing was used to detect the variations in HBV quasispecies. The concentrations of the PCR products were measured using the Qubit high-sensitivity double-stranded DNA assay (Invitrogen, Carlsbad, CA, USA). Next, a library of the viral genomic PCR products (< 500 bp; 50 or 0.2 ng) was prepared using the Nextera DNA sample prep kit (Illumina, San Diego, CA, USA) or Nextera XT DNA sample prep kit (Illumina), according to the manufacturer's instructions. The PCR products were uniformly sheared to 500 bp fragments using these kits, and the PCR products were mixed with 1% 8 pM PhiX as a control and run on the MiSeq sequencer (Genome Analyzer; Illumina) for paired-end 151 bp sequencing. Finally, image-based detection of the fluorescent signals was performed using MiSeq control software (Illumina), and the images were used to obtain the sequence data in FASTQ format. Illumina paired-end sequencing was performed to generate overlapping read pairs among relatively short sequences combined with relatively long sequences. A quality check and data trimming were necessary before assembling the sequences using Genomics Workbench software version 6.0.1 (CLC bio, Aarhus, Denmark). The sequencing results used in this analysis had a read quality of > 80% of the consensus sequence estimated from a quality score of 30 according to the manufacturer's (Illumina) data. Overlapping read pairs and filtering of reads based on a quality score of 30 were used to reveal sequencing errors and eliminate any false-positive variants generated by PCR errors during sequencing, to increase

Hatazawa *et al. Infectious Agents and Cancer* (2018) 13:7

Page 3 of 8

the confidence that the sequence reads can detect viral variants of low abundance.

This study was approved by the Ethical Review Board of Kobe University.

Statistical analysis

Statistical analyses were performed using SPSS version 24.0 (IBM Corporation, S&I, Tokyo, Japan). Differences between groups were examined by Student's *t*-test or analysis of variance. *P* values < 0.05 were considered statistically significant.

Results

Clinical characteristics

The average age of the OBI-related HCC patients was significantly higher than that of the HBsAg-positive HCC patients (67.7 ± 5.6 vs. 53.6 ± 11.5 years, respectively). There was no statistical significance of platelet count and transaminase levels between two groups. Among the OBI-related HCC patients, serum HBV DNA levels were undetectable, but the anti-HBc titre was 10.1 ± 1.2 (range 8.56–12.27) S/CO. Regarding the staging of liver fibrosis, no difference was found in the hyaluronic acid level and the proportion of histologically F3 (pre-cirrhosis) and F4 (cirrhosis) patients at the time of surgery between the groups (OBI 56% vs. HBsAg-positive 60%) (Table 1).

Ultra-deep sequencing analysis

The numbers of mapped reads and average coverage of the pre-S/S region were $859,778 \pm 682,646$ (range 270,715–2,941,784) and $106,038 \pm 81,043$ (range 30,487–301,730), respectively. The amino acids (aa) accounting

for more than 1% of the variation comprised 55.1% of the aa in the pre-S/S domain, with no significant difference between the OBI and HBsAg-positive groups (Table 2).

The consensus sequence was determined by the NCBI Blast program (available at https://blast.ncbi.nlm.nih.gov/Blast.cgi) using the sequences obtained in this study. Based on the phylogenetic analysis, all of our strains were classified into genotype C2. The closest reference HBV genotype was C2 (GenBank accession no. AB014394). The viral quasispecies were evaluated in terms of their prevalence in the viral population using the setting "read conflict" in Genomics Workbench. Based on the nucleotide alterations detected, the viruses with aa changes were determined to be variants. The proportion of a variant population was determined as the percentage of aa substitutions per aa coverage depth (entire number of sequence reads at each position) [19].

The aim of this study was to reflect the quasispecies variants in relation to hepatocarcinogenesis; therefore, we analysed variants detected in > 1% of the total viral population. A similar cut-off was applied in previous studies [20, 21]. In addition, > 20% of the total viral population is possible to detect the conventional sequencing method. In this reason, the viral quasispecies were divided and analysed; viral quasispecies detected in > 1% of the total viral population were classified as a major population (20–80% of the total population) and a minor population was defined as that comprising > 1% to < 20% of the total population, respectively.

Amino acid variations in the pre-S region

Within the 174 aa pre-S region, the number of major variant populations detected were 4.22 ± 3.31 in OBI-

Table 1 Clinical and virological characteristics of the patients

	HBsAg-positive HCC ($n = 10$)	OBI-related HCC ($n = 9$)	*P* value
Age (years)	53.6 ± 11.5	69.7 ± 5.6	0.002*
Sex (male/female)	9/1	8/1	0.737
PLT ($\times 10^4$/mm3)	17.9 ± 4.9	23.0 ± 8.4	0.131
AST (U/L)	59.1 ± 37.9	38.9 ± 15.0	0.145
ALT (U/L)	33.2 ± 11.0	31.4 ± 15.4	0.782
HbA1c (%)	5.29 ± 0.45	5.46 ± 0.50	0.461
HBsAg (mIU/ml)	3278 ± 3045	0	
Anti-HBc antibody (C.O.I.)	12.6 ± 1.8	10.1 ± 1.2	0.002*
HBV DNA (log copies/ml)	5.41 ± 1.84	< 2.06	
Hyaluronic acid (ng/ml)	62.7 ± 51.0	107.1 ± 123.2	0.338
Fibrosis (F0–2/F3–4)	4/6	4/5	0.605
Total reads (mean ± SD)	$4,451,542 \pm 2,357,824$	$3,708,797 \pm 2,365,839$	0.503
Mapping reads (mean ± SD)	$1,075,698 \pm 845,602$	$619,866 \pm 347,934$	0.144
Coverage (mean ± SD)	$131,921 \pm 97,014$	$77,279 \pm 49,215$	0.139

*$P < 0.05$

Hatazawa *et al. Infectious Agents and Cancer* (2018) 13:7

Page 4 of 8

Table 2 Amino acid mutations in the pre-S/S region of the major quasispecies populations

Domain	AA Position (number)	HBsAg Positive (%)	OBI (%)	P
PreS1	aa1–119 (119)	2.80 ± 1.69	4.22 ± 3.31	0.12
NTCP	aa2–48 (47)	0.60 ± 0.84	0.33 ± 0.50	0.21
HSP70	aa74–118 (45)	0.60 ± 0.70	0.33 ± 0.50	0.18
S promoter	aa66–111 (46)	0.70 ± 0.67	0.33 ± 0.50	0.10
NBS	aa103–127 (25)	0.20 ± 0.63	0.56 ± 0.88	0.16
T cell epitope	aa21–30,52–67 (26)	0.50 ± 0.71	0.67 ± 0.87	0.33
B cell epitope	aa12–47,72–78,94–117 (67)	0.70 ± 0.82	0.00 ± 0.00	0.01*
S	aa1-227 (227)	8.10 ± 8.62	7.33 ± 4.74	0.41
Pre-a determinant	aa100–119 (20)	1.20 ± 0.79	1.11 ± 0.33	0.38
a determinant	aa 124–148 (25)	0.70 ± 0.82	0.00 ± 0.00	0.01*
Post a determinant	aa 149–169 (21)	0.30 ± 0.67	0.22 ± 0.67	0.40
HLA class I	aa87–98, 186–197, 215–223 (29)	1.80 ± 1.81	1.44 ± 0.73	0.29
HLA class II	aa97–106, 171–179, 206–215 (29)	0.50 ± 0.85	0.44 ± 0.73	0.44

*P < 0.05

related HCC and 2.80 ± 1.69 in HBsAg-positive HCC, respectively. Although no statistical difference in the proportion of the major quasispecies populations in the entire pre-S or pre-S1 regions was detected, the proportion in the pre-S2 region was significantly higher in OBI-related HCC than in HBsAg-positive HCC (2.89 ± 2.32% vs. 1.30 ± 0.95%, P = 0.041) (Fig. 1).

According to the various domains of the pre-S region, major HBV populations based on aa mutations in the B-cell epitope within the pre-S1 region were detected only in HBsAg-positive HCC (P = 0.01) (Table 2). On the other hand, the number of minor populations based on mutations in the polymerized human serum albumin (pHSA) domain within the pre-S2 region was significantly higher in OBI-related HCC compared with HBsAg-positive HCC (2.44 ± 2.30 vs. 0.80 ± 1.48, P = 0.039). The different mutations detected were W4R, S5A/P/T, L30S, G35R, H51P/R, E54K, A60V, W77R/Stop, I84T, N98 K/T, G102R, Q118R/Stop, N123D and S124F/P. Among them, the W4R, L30S and Q118R/Stop mutations were more prevalent in OBI-related HCC (Table 3).

Amino acid variations in the S region

In the 227 aa S domain, major HBV quasispecies populations were detected in 7.33 ± 4.74 of OBI-related HCC and 8.10 ± 8.62 of HBsAg-positive HCC, but there was no significant difference between the groups (Fig. 1). Major populations based on the alpha determinant region (aa 124–148, 25 aa) were detected in HBsAg-positive HCC (P = 0.01) but not in OBI-related HCC (Table 2). Differences in the major populations were attributed to the I126S/T, Y200F/S, and Y206C/H/S mutations, while those in the minor populations were

attributed to the L21F/S, L42F/S, W182Stop, and L213F/I/T mutations. I126S/T, known as an escape mutation in the alpha determinant region, was found only in HBsAg-positive HCC (Table 3).

Discussion

Several epidemiological and molecular studies have reported OBI as a risk factor for cirrhosis and HCC [10, 22–25]. The detection of OBI is variable in depend on the HBsAg detection kit. In particular, several older HBsAg detection kits cannot detect escape mutations in the alpha determinant region. More recently, improvements in HBsAg detection kits have made it possible to detect HBsAg, including escape mutants. The sensitivity of HBsAg detection is also improving, and most of the available kits can detect very low levels of HBsAg. The Lumipulse HBsAg HQ assay, which was used in this study, can detect extremely low levels of HBsAg, with a cut-off level of > 0.0005 IU/ml [26]. In this study, all serum HBV DNA levels were below the sensitivity threshold but HBV DNA was detected in the liver tissues. OBI was clearly confirmed, even when using a highly sensitive HBsAg kit.

All patients who were negative for HBsAg but positive for anti-HBc were treated in this study. This condition is generally considered seropositive OBI, and is treated separately from anti-HBc- and anti-HBs-seronegative OBI [9, 10]. Seropositive OBI is considered to indicate the resolved phase of acute hepatitis or the recovery phase of chronic hepatitis. Chu and Liaw reported that the presence of cirrhosis is an independent predictive factor for HBsAg seroclearance [4]. In this study, all patients showed high anti-HBc antibody titres and were thought to have had overt chronic active hepatitis. In

Hatazawa *et al. Infectious Agents and Cancer* (2018) 13:7

Page 5 of 8
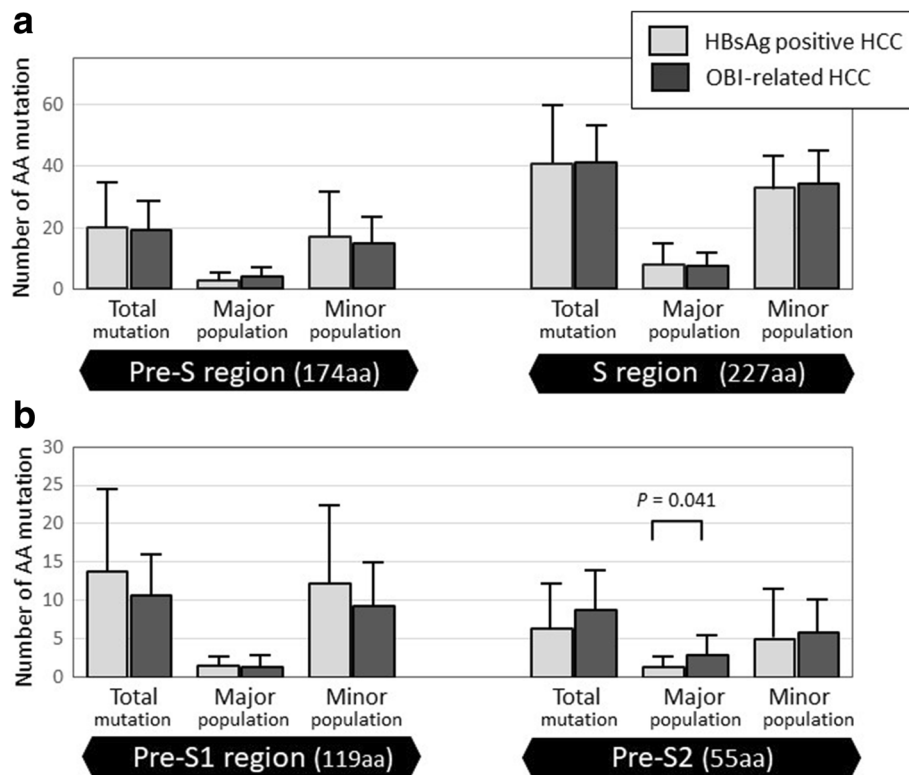
Fig. 1 Number of aa mutations in the pre-S/S region. Number of aa mutations in the pre-S/S region (**a**), and mutations were also shown in pre-S1 and pre-S2 region (**b**). No significant difference in the average number of aa mutations was detected between HBsAg-positive and OBI-related HCC. The proportion of the major variant populations based on the pre-S2 region was significantly higher in OBI-related HCC than in HBsAg-positive HCC

fact, there was no difference in liver fibrosis between the OBI and HBsAg-positive cases, and five of the nine OBI cases had a histologically advanced stage (F3/4).

HBV generally exists as numerous quasispecies in the host because of the lack of a proofreading function during the reproduction process. Many studies have reported that the small number of viral variants was related to disease progression and the therapeutic response. It is thought that the small number of viral variants and the extent of variation are partly related to the clinical condition. Several reports revealed that the variation in, or deletion of, the pre-S/S region was related to the development of liver cancer as well as disappearance of the HBsAg in OBI patients. Previously, the M1L, Q2K and S182Stop in the pre-S2 region and T105C in the pre-S/S region were reported as specific mutations in OBI patients [27]. In this study, we could not detect any unique mutations in the OBI patients. So far, direct sequencing is the most common method used to detect viral mutations. However, variations present in approximately less than 20% of viral quasispecies cannot be detected by direct sequencing [28, 29]. The recent introduction of next-generation sequencing enabled analysis of viral sequence variations in more detail. Using next-generation sequencing, an increased number of reports have revealed that minor variants are related to the clinical condition [30, 31]. In this study, several mutations in the minor populations were detected in OBI patients. Further studies are necessary to clarify the clinical implications.

The HBsAg is produced from the S region of HBV. The pre-S/S region contains many important domains, including the major hydrophilic region (MHR) related to the host immune response, a specific domain recognizing hepatocytes in the host, and an antigen determinant region (Fig. 2). The HBsAg becomes the target of host immunoreactivity by vaccine-induced and naturally acquired antibodies; in particular, the domain of the aa 100–169 is known as MHR, which contains the B-cell epitope and is recognized by HBs antibody. Though the detailed border of the recognition part is unknown, it has been clarified that antigen determining region α (aa124–147) is strongly related to the binding by anti-HBs antibody. In addition, it is known that the mutations and the variety in MHR domain are related to the antigenic change of the HBsAg, the evasion of the immune response from HBs antibody, the responsiveness from immune induction therapy such as globulin, and a
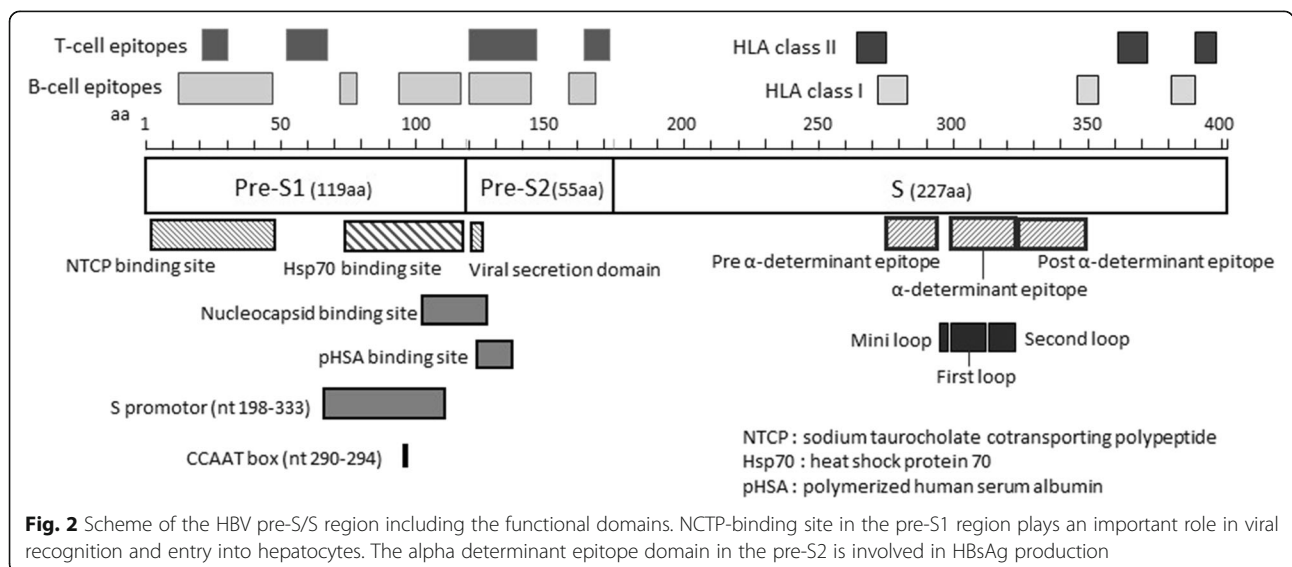
Hatazawa et al. Infectious Agents and Cancer (2018) 13:7

Page 6 of 8

Table 3 The frequencies of specific amino acid mutations associated with the major and minor population of HBV quasispecies

Region	Position	HBsAg positive (n = 10)	OBI (n = 9)	P value
Major population				
S	I126S/T	30	0	0.040
	Y200F/S	**0**	**33.3**	**0.025**
	Y206C/H/S	30	0	0.040
Minor population				
Pre-S	**W4R**	**0**	**33.3**	**0.025**
	S5A/P/T	30	0	0.040
	L30S	**10**	**77.8**	**0.001**
	G35R	30	0	0.040
	H51P/R	30	0	0.040
	E54K	50	0	0.006
	A60V	30	0	0.040
	W77R/Stop	30	0	0.040
	I84T	60	11.1	0.014
	N98 K/T	30	0	0.040
	G102R	30	0	0.040
	Q118R/Stop	**0**	**44.4**	**0.008**
	N123D	**10**	**44.4**	**0.049**
	S124F/P	**0**	**33.3**	**0.025**
S	L21F/S	70	100	0.040
	L42F/S	**0**	**33.3**	**0.025**
	W182Stop	30	0	0.040
	L213F/I/T	40	0	0.017

Bold indicated the amino acid position which mutational frequencies in OBI were higher than those in HBsAg positive

diagnostic failure by the conventional diagnosis tool of HBsAg [32, 33]. In this study, the prevalence of the major variant population in antigen determining region α was significantly lower among OBI-related HCC than in HBsAg-positive HCC. We suggest that this region would be conserved in OBI patients because the host immune response was smaller than that in HBsAg-positive patients. Several specific mutations, including G119R, C124Y, I126S, Q129R, S136P, C139R, T140I, K141E, D144A, and G145R, have been reported as being related to the decrease of HBsAg secretion [34]. In present study, I126S/T mutants were detected only in HBsAg-positive cases, and it was suggested that the titre of HBsAg of these cases might decrease over a long-term follow-up period.

The present study revealed that the prevalence of the major variant populations associated with the B-cell epitope, within the pre-S1 region, was significantly higher in HBsAg-positive HCC than in OBI-related HCC. The greater variation in the pre-S/S region, especially the B-cell epitope, in HBsAg-positive HCC compared with OBI-related HCC, is possibly a mechanism to avoid immune pressure from the host. In contrast, the prevalence of major populations in the pre-S2 region, especially the pHSA domain, was significantly higher in OBI-related HCC than in HBsAg-positive HCC. The pre-S2 region comprises 55 aa and, together with the pre-S1 and S regions, encodes the surface proteins of HBV [35, 36]. In addition, the pre-S2 region contains the T-cell epitope, B-cell epitope, and pHSA-binding domain domains (Fig. 2). pHSA is one of the receptors that binds the HBV Dane particle [37]. The pre-S2 region is associated with HCC. It was reported that deletion of aa 1–6 of the pre-S2 region, encoding the B cell epitope, is related to HCC development [38]. In addition, a previous study reported

Fig. 2 Scheme of the HBV pre-S/S region including the functional domains. NCTP-binding site in the pre-S1 region plays an important role in viral recognition and entry into hepatocytes. The alpha determinant epitope domain in the pre-S2 is involved in HBsAg production

Hatazawa et al. Infectious Agents and Cancer (2018) 13:7

Page 7 of 8

that aa mutations in the pre-S2 region, especially the pHSA domain, are prevalent in HBV-related HCC [36], and HBV-infected patients with mutations in the pHSA region frequently progress to cirrhosis and HCC [39]. These previous studies suggested that the pre-S2 region does not play an important role in the synthesis or secretion of viral particles or in the infectiousness of the virus, but the accumulation of viral particles in hepatocytes with pHSA mutations is related to cytotoxic accumulation and progression to cirrhosis and HCC. The results of the present study suggest that a high level of variation in the pre-S2 region, particularly the pHSA domain, is related to hepatocarcinogenesis, even in OBI patients with low viral tites.

The limitations of the present study include the small number of Japanese cases evaluated. In addition, the average age of present study is different in two groups. It is usual that HBsAg positive HCC patients are younger than other cause of HCC patients. Although HBV is naturally mutated over time, it was none the less important that variation in the pre-S2 was detected in OBI patients. Future studies with larger cohorts including several viral genotypes and various host genetic factors are required to overcome these limitations.

Conclusion

Viral quasispecies based on the pre-S/S region were less diverse in OBI-related HCC than in HBsAg-positive HCC. The most significant variation was in the B cell epitope within the pre-S1 region and the MHR domain within the S region. It is suggested that the evasion mechanism from the host immune response is weaker in OBI-related HCC. Furthermore, because mutations in the pre-S2 region were significantly prevalent in OBI-related HCC, variation in this region may be related to hepatocarcinogenesis in OBI patients.

Abbreviations
HBV: Hepatitis B virus; HCC: Hepatocellular carcinoma; OBI: Occult HBV infection

Availability of data and materials
The data sets supporting the conclusions of this article are included within the article.

Authors' contributions
YH, YY, and YH designed the study; MT, and TF were responsible for the recruitment of subjects; RO, TT, HH, HH, AM, YK, YM, and MY performed experiments and conducted data management; YH and YY wrote the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate
The study protocol conformed to the ethical guidelines of the 1975 Declaration of Helsinki and was approved by the Medical and Health Research Ethics Committee (MHREC), Kobe University Graduate School of Medicine (Trial registration: Ref: 1856 [Approved 22 March 2016]). All subjects gave their written informed consent before enrolment.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
[1]Division of Gastroenterology, Department of Internal Medicine, Kobe University Graduate School of Medicine, Kobe, Hyogo 650-0017, Japan. [2]Division of Molecular Medicine & Medical Genetics, Department of Pathology, Kobe University Graduate School of Medicine, Kobe, Japan. [3]Department of Internal Medicine, Tokushima Prefectural Naruto Hospital, Tokushima, Japan. [4]Division of Hepato-Biliary-Pancreatic Surgery, Department of Surgery, Kobe University Graduate School of Medicine, Kobe, Japan. [5]Department of Hepatology, Osaka City University Graduate School of Medicine, Osaka, Japan.

References
1. Schweitzer A, Horn J, Mikolajczyk RT, Krause G, Ott JJ. Estimations of worldwide prevalence of chronic hepatitis B virus infection: a systematic review of data published between 1965 and 2013. Lancet. 2015;386:1546–55.
2. Lee WM. Hepatitis B virus infection. N Engl J Med 1997; 337: 1733–1745.
3. Hoofnagle JH, Doo E, Liang TJ, Fleischer Rand Lok AS. Management of hepatitis B: summary of a clinical research workshop. Hepatology. 2007;45:1056–75.
4. Chu CM, Liaw YF. HBsAg seroclearance in asymptomatic carriers of high endemic areas: appreciably high rates during a long-term follow-up. Hepatology. 2007;45:1187–92.
5. Wong RJ, Nguyen MT, Trinh HN, Chan C, Huynh A, Ly MT. et al, Hepatitis B Surface Antigen Loss and Sustained Viral Suppression in Asian Chronic Hepatitis B Patients: A Community-Based Real World Study. J Viral Hepat. 2017 Jun 5; doi: 10.1111/jvh.12736.
6. Liu J, Liu J, Yang HI, Lee MH, Lu SN, Jen CL, et al. REVEAL-HBV Study Group . Incidence and determinants of spontaneous hepatitis B surface antigen seroclearance: a community-based follow-up study. Gastroenterology. 2010; 139:474–82.
7. European Association for the Study of the Liver: European Association for the Study of the Liver. EASL 2017 Clinical Practice Guidelines on the management of hepatitis B virus infection. J Hepatol. 2017; 67: 370–98.
8. Tuttleman JS, Pourcel C, Summers J. Formation of the pool of colalently closed circular viral DNA in hepadnavirus-infected cells. Cell. 1986;47:451–60.
9. Raimondo G, Allain JP, Brunetto MR, Buendia MA, Chen DS, Colombo M, et al. Statements from the Taormina expert meeting on occult epatiti B virus infection. J Hepatol. 2008;49:652–7.
10. Raimondo G, Caccamo G, Filomia R, Pollicino T. Occult HBV infection. Semin Immunopathol. 2013;35:39–52.
11. Huang X, Qin Y, Li W, Shi Q, Xue Y, Li J, et al. Molecular analysis of the hepatitis B virus presurface and surface gene in patients from eastern China with occult hepatitis B. J Med Virol. 2013;85:979–86.
12. 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.
13. Sinn DH, Choi MS, Gwak GY, Paik YH, Lee JH, Koh KC, et al. Pre-S mutation is a significant risk factor for hepatocellular carcinoma development: a long-term retrospective cohort study. Dig Dis Sci. 2013;58:751–8.

Hatazawa *et al. Infectious Agents and Cancer* (2018) 13:7

Page 8 of 8

14. Rodriguez C, Chevaliez S, Bensadoun P, Pawlotsky JM. Characterization of the dynamics of hepatitis B virus resistance to adefovir by ultra-deep sequencing. Hepatology. 2013;58:890–901.

15. Mese S, Arikan M, Cakiris A, Abaci N, Gumus E, Kursun O, et al. Role of the line probe assay INNO-LiPA HBV DR and ultradeep sequencing in detecting resistance mutations to nucleoside/nucleotide analogues in viral samples isolated from chronic hepatitis B patients. J Gen Virol. 2013;94:2729–38.

16. Zhang XX, Li MR, Cao Y, Zhang RW, Zhang Y, Li F, et al. Dynamics of genotypic mutations of the hepatitis B virus associated with long-term Entecavir treatment determined with Ultradeep sequencing: a retrospective observational study. Medicine (Baltimore). 2016;95:e2614.

17. Yin F, Wu Z, Fang W, Wu C, Rayner S, Han M, et al. Resistant mutations and quasispecies complexity of hepatitis B virus during telbivudine treatment. J Gen Virol. 2015;96:3302–12.

18. Sugauchi F, Mizokami M, Orito E, Ohno T, Kato H, Suzuki S, et al. A novel variant genotype C of hepatitis B virus identified in isolates from Australian aborigines: complete genome sequence and phylogenetic relatedness. J Gen Virol. 2001;82:883–92.

19. Yamani LN, Yano Y, Utsumi T, Juniastuti WH, Widjanarko D, et al. Ultradeep sequencing for detection of quasispecies variants in the major hydrophilic region of hepatitis b virus in Indonesian patients. J Clin Microbiol. 2015;53:3165–75.

20. Bolcic F, Sede M, Moretti F, Westergaard G, Vazquez M, Laufer N, Quarleri J. Analysis of the PKR-eIF2 alpha phosphorylation homology domain (PePHD) of hepatitis C virus genotype 1 in HIV-coinfected patients by ultra-deep pyrosequencing and its relationship to responses to pegylated interferon-ribavirin treatment. Arch Virol. 2012;157:703–11.

21. Eriksson N, Pachter L, Mitsuya Y, Rhee SY, Wang C, Gharizadeh B, Ronaghi M, Shafer RW, Beerenwinkel N. Viral population estimation using pyrosequencing. PLoS Comput Biol. 2008;4:e1000074.

22. Hass M, Hannoun C, Kalinina T, Sommer G, Manegold C, G€unther S. Functional analysis of hepatitis B virus reactivating in hepatitis B surface antigen-negative individuals. Hepatology. 2005;42:93–103.

23. Mrani S, Chemin I, Menouar K, Guillaud O, Pradat P, Borghi G, et al. Occult HBV infection may represent a major risk factor of non-response to antiviral therapy of chronic hepatitis C. J Med Virol. 2007;79:1075–81.

24. De Mitri MS, Cassini R, Bernardi M. Hepatitis B virus-related hepatocarcinogenesis: molecular oncogenic potential of clear or occult infections. Eur J Cancer. 2010;46:2178–86.

25. Chemin I, Zoulim F, Merle P, Arkhis A, Chevallier M, Kay A, et al. High incidence of hepatitis B infections among chronic hepatitis cases of unknown etiology. J Hepatol. 2001;34:447–54.

26. Zacher BJ, Moriconi F, Bowden S, Hammond R, Louisirirotchanakul S, Phisalprapa P, et al. Multicenter evaluation of the Elecsys hepatitis B surface antigen quantitative assay. Clin Vaccine Immunol. 2011;18:1943–50.

27. Coppola N, Onorato L, Iodice V, Starace M, Minichini C, Farella N, et al. Occult HBV infection in HCC and cirrhotic tissue of HBsAg negative patients: a virological and clinical study. Oncotarget. 2016;7:62706–14.

28. Hussain M, Fung S, Libbrecht E, Sablon E, Cursaro C, Andreone P, et al. Sensitive line probe assay that simultaneously detects mutations conveying resistance to lamivudine and adefovir. J Clin Microbiol. 2006;44:1094–7.

29. Lok AS, McMahon BJ. Chronic hepatitis B. Hepatology. 2007;45:507–39.

30. Suzuki F, Hosaka T, Suzuki Y, Akuta N, Sezaki H, Hara T, et al. Long-term efficacy and emergence of multidrug resistance in patients with lamivudine-refractory chronic hepatitis B treated by combination therapy with adefovir plus lamivudine. J Gastroenterol. 2014;49:1094–104.

31. Wang YW, Shan X, Huang Y, Deng H, Huang WX, Zhang DZ, et al. A novel baseline hepatitis B virus sequencing-based strategy for predicting adefovir antiviral response. Infect Genet Evol. 2015;33:269–76.

32. Kay A, Zoulim F. Hepatitis B virus genetic variability and evolution. Virus Res. 2007;127:164–76.

33. Ogura Y, Kurosaki M, Asahina Y, Enomoto N, Marumo F, Sato C. Prevalence and significance of naturally occurring mutations in the surface and polymerase genes of hepatitis B virus. J Infect Dis. 1999;180:1444–51.

34. Huang CH, Yuan Q, Chen PJ, Zhang YL, Chen CR, Zheng QB, et al. Influence of mutations in hepatitis B virus surface protein on viral antigenicity and phenotype in occult HBV strains from blood donors. J Hepatol. 2012;57:720–9.

35. Locarnini S, McMillan J, Bartholomeusz A. The hepatitis B virus and common mutants. Semin Liver Dis. 2003;23:5–20.

36. 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;130:1153–68.

37. Okamoto H, Usuda S, Imai M, Tachibana K, Tanaka E, Kumakura T, et al. Antibody to the receptor for polymerized human serum albumin in acute and persistent infection with hepatitis B virus. Hepatology. 1986;6:354–9.

38. Kao JH, Liu CJ, Jow GM, Chen PJ, Chen DS, Chen BF. Fine mapping of hepatitis B virus pre-S and its association with hepatocellular carcinoma. Liver Int. 2012;32:1373–81.

39. Gao ZY, Li T, Wang J, Du JM, Li YJ, Li J, Lu FM, Zhuang H. Mutations in preS genes of genotype C hepatitis B virus in patients with chronic hepatitis B and hepatocellular carcinoma. J Gastroenterol. 2007;42:761–8.