Hepatitis B Virus Core Protein Variations Differ in Tumor and Adjacent Nontumor Tissues from Patients with Hepatocellular Carcinoma

Yu Zhu, Yan Jin, Xiaojin Cai, Xin Bai, Min Chen, Taoyang Chen, Jinbing Wang, Gensun Qian, Jianren Gu, Jinjun Li, Hong Tu

Objectives: To characterize the mutation pattern of a hepatitis B virus (HBV) core protein (HBcAg) derived from hepatocellular carcinoma (HCC) and adjacent nontumor tissues.

Methods: HBV core gene fragments (nt. 1901–2365) were amplified from 98 HBV-related HCC tissues and 33 adjacent nontumor tissues. The deduced amino acids (AAs) of the core gene were aligned with the prototype sequences of HBV genotypes B and C.

Results: In total, there were 54 positions that showed polymorphism at the deduced AA level. The mutations were predominantly located in three major (codons 83–87, 95–104 and 130–135) and three minor (codons 21–38, 59–63 and 151–155) mutation-clustering regions (MCRs). The substitution rate in MCRs was significantly higher than in mutation-devoid regions (p < 0.001). The most frequently occurring mutations in rank were codon P130T (38.8%), I97L (37.8%) and S87G (23.5%). In addition, there were 7 patients that showed internal deletions in the middle of HBcAg with sizes ranging from 34 to 59 AAs. Unexpectedly, the core genes isolated from tumor tissues had fewer mutations compared with those isolated from adjacent nontumor tissues from the same patients (p < 0.05).

Conclusions: Accumulation of naturally occurring mutations in certain restricted segments of HBcAg may be related to the development of HCC.

Key Words
Core protein • Hepatitis B virus • Hepatocellular carcinoma • Mutation-clustering region

Introduction

Worldwide, more than 350 million people are chronically infected with the hepatitis B virus (HBV) and are at increased risk of developing hepatocellular carcinoma (HCC) [1]. In addition to host factors, viral factors per se can also play an important role in determining clinical outcomes [2, 3].

HBV is a small, enveloped DNA virus of the Hepadnaviridae family. The virus has a partially double-stranded DNA genome of about 3.2 kb with four overlapping open reading frames that encode the viral envelope protein (pre-S/S), the X protein (X), the DNA polymerase (P) and the core protein (precore/core). HBV replicates via a reverse transcription reaction with an RNA intermediate. As the reverse transcriptase lacks proofreading activity, errors in HBV replication occur at a much higher rate than in other DNA viruses. Hence, various mutations can be observed in the HBV genome during long-term infec-
tion; some of these markers can serve as viral markers for predicting the development of HBV-associated HCC. The HBV genotype, basal core promoter mutation, pre-core mutation and pre-S deletion have been implicated in the development of HCC [4–8]. However, the association of core gene mutation and HCC has been largely unexplored. While several mutation hot-spots, such as codons 5, 13, 59, 60, 87, 97 and 130 of the hepatitis B core antigen (HBcAg), have been identified in HBV chronic carriers [9–13], only a few have been reported to increase the risk of HCC [14]. Recently, six mutations at HBcAg codons 13, 21, 49, 79, 113 and 181 were reported to decrease the risk of HCC. The protective effects of these variants on HCC may be attributable to the reduced level of viral load and HBV e antigen negativity [15].

Qidong, an east coastal city in Jiangsu province, China, is a hyperendemic area of HBV. The province also leads in the incidence of HCC in China [16, 17]. To assess the potential role of HBV core gene mutations in developing HCC, we examined the HBV sequences from nt. 1901 to 2365 in 98 HCC tissues collected in Qidong City. Thirty-three paired tumor and adjacent nontumor tissues were employed, for the first time, to compare the mutation rate and pattern of HBcAg in these tissues.

Materials and Methods

Patients and Tissue Samples

Ninety-eight HCC tumor tissues and 33 corresponding nontumor liver tissues were obtained from patients with HCC who underwent surgical resection at the Qidong Liver Cancer Institute/Qidong Tumor Hospital from 1996 to 2006. All participants were histologically diagnosed with HCC and were seropositive for hepatitis B surface antigen and HBV DNA (mean ± SD, 3.68 ± 0.81 log copies/ml). HCC patients were at stage A1 (BCLC staging system). Tumor grade was II, II–III and III in 47, 9 and 42 cases, respectively. Patients with hepatitis C virus coinfection were excluded from the study. Ninety patients were men and 8 were women. Patients' ages ranged from 32 to 75 years (mean, 49 years). The average duration of hepatitis B before the diagnosis of HCC was 14.5 years. None of the patients had received treatment with interferon or nucleotide analogues. Thirty patients had been treated with Chinese herbal preparations to alleviate the inflammation and improve liver function. There were 55 patients who had been diagnosed with liver cirrhosis before HCC occurrence. Fresh tumor and nontumor liver tissues were snap-frozen in liquid nitrogen and stored at −70°C until use. This work was performed in accordance with the Helsinki Declaration. Written informed consent was obtained from all patients and the study protocol was approved by the institutional review boards of the Qidong Liver Cancer Institute/Qidong Tumor Hospital and the Shanghai Cancer Institute, Renji Hospital, Shanghai Jiao Tong University School of Medicine.

DNA Extraction from Tissue

One hundred milligrams from each sample of frozen liver tissue was cut into small fragments and incubated with 600 μl lysis buffer (10 mM Tris-HCl, pH 8.0, 10 mM EDTA, 2.5% SDS, 200 μg/ml proteinase K) at 50°C for 20 h. After the tissues were homogenized, genome DNA was extracted by the conventional phenol-chloroform method.

Amplification and Sequencing of HBV Core Gene

The HBV core gene from nt. 1901 to 2365 was amplified by semi-nested PCR using the sense primer pre-C F1, 5′-TTCCACCTCTGCCTAACATCATCTC-3′ (nt. 1824–1845) and the antisense primer HBV2570R, 5′-AATTGCCTCTTGTAATTGATG-3′ (nt. 2570–2549) as the first round primers, and the sense primer pre-C F1 and the antisense primer HBV2433R, 5′-GATTGAGATCTTCTGCGACGC-3′ (nt. 2433–2413) as the second round primers. PCR was performed under the following conditions: 94°C for 3 min, followed by 94°C for 30 s, 58°C for 30 s and 72°C for 1 min for 35 cycles with a final extension at 72°C for 7 min. The PCR products were gel purified (Axygen Scientific, Inc., Union City, Calif., USA) and used as templates for sequencing using the BigDye terminator cycle–sequencing reaction kit and the Prism 3700 DNA analyzer (Applied Biosystems, Foster City, Calif., USA). Sequences of the core gene were compared using MEGA4.1 [18].

HBV Genotyping

HBV genotypes were determined by comparing the sequence of the X gene with a set of database-derived standard sequences. Standard sequences were retrieved from GenBank/DDBJ/EMBL. The phylogenetic tree was constructed with the software MEGA4.1 [18].

Statistical Analysis

Student's t test was used for continuous variables with normal distributions. All of the tests were two-tailed, and a p value of <0.05 was considered statistically significant. SPSS (SPSS Inc. Chicago, Ill., USA) version 12.0 was used for the statistical analysis.

Results

Nucleotide Sequence and Deduced Amino Acid Residue of the HBV Core Gene Isolated from 98 HCC Tissues

HBV DNA fragments (nt. 1901–2365) were successfully amplified by PCR from 98 HCC tissues. This region was translated into the first 155 amino acids (AAs) of the core protein. Among the 98 HBV isolates, 90 belonged to genotype C and 8 belonged to genotype B. Available nucleotide sequences GU434373 and GU434374 from GenBank were used as the prototype sequences for genotype B and C, respectively. These two viral strains were isolated from HBV carriers in the Qidong area. In total, 54 sites with missense mutations were found when aligned
to the respective prototype sequence (fig. 1). Seventy-nine patients (80.6%) showed at least one missense mutation within the detected core region. The mutations most frequently occurred at codon 130 (38.8%), codon 97 (37.8%) and codon 87 (23.5%). The substitution rate at each AA residue was 0.0123 ± 0.0333.

**Mutation-Clustering Regions and Hot-Spot Sites of HBV Core Protein**

The distribution pattern of 54 AA alterations is presented in figure 1. Most of the changes clustered in small segments. The six mutation-clustering regions (MCRs) were codons 21–38 (MCR1), codons 59–63 (MCR2), codons 83–87 (MCR3), codons 95–104 (MCR4), codons 130–135 (MCR5) and codons 151–155 (MCR6). We tentatively defined MCR3, MCR4 and MCR5 as the major domains and MCR1, MCR2 and MCR6 as the minor domains. The mutation rates ranged from 0.0561 to 0.0694 in the major domains and from 0.0249 to 0.0421 in the minor domains. In addition to the MCRs, two point mutations located at codon 5 and codon 13 were found to occur with high frequency. Between the MCRs and hot-spot point mutations, eight mutation-devoid regions (MDRs) were tentatively defined. The average mutation rate in the MDRs was 0.0032 ± 0.0078, which was significantly lower than that in the major domains (p < 0.001) or minor domains (p < 0.001). These data indicate that the distribution of mutations in HBV core protein is not random.

**Internal Deletion of HBV Core Protein**

Among the 98 HBV sequences isolated from HCC tissues, 7 (7.1%) had internal deletions in the core gene. The deletions ranged in size from 102 to 177 bp, which caused 34–59 AA truncations in the middle of the core protein (fig. 2). No deletion caused frameshift alterations as the

---

**Fig. 1.** Distribution of AA changes in the HBV core protein. The first 155 AAs of the core protein were aligned with prototype sequences of HBV. Six MCRs and eight MDRs were identified. WT = Wild type.
sizes of the nucleotide deletion were in multiples of three. In all 7 cases, the deletions terminated upstream of the AUG translation initiation codon of the P gene (nt. 2307–2309, corresponding to the core gene codon 136), thus leaving the HBV polymerase intact. The common region affected by the truncations was from codon 84 to 97. If HCC case T98 were not considered, the common truncated region would extend up to codon 110. This region (codons 84–110) involved two major mutation clustering domains, namely MCR3 and MCR4.

Comparison of Core Mutations between 33 Paired Tumor and Nontumor Tissues

Thirty-three HCC tissues and the corresponding adjacent nontumor liver tissues were included in the study to compare the mutation rate and pattern of the HBV core gene. No internal deletion existed in these 33 paired tumor and nontumor tissues. The mutation rate of HBcAg in nontumor tissues was significantly higher than that in tumor tissues (0.0217 ± 0.0163 vs. 0.0141 ± 0.0112, p < 0.05). Further analysis revealed that the elevated rate of mutation in nontumor tissues was mainly attributed to the AA changes in a small region between codon 39 and 82. This region is composed of one MCR (MCR 2) and two MDRs (MDR4/MDR5). While core sequences from 33 tumor tissues showed no missense mutation in MDR4 and MDR5, those from 33 nontumor tissues displayed the missense mutations at a rate of 0.0091 ± 0.0232 for MDR4 (p < 0.05) and 0.0207 ± 0.0415 for MDR5 (p < 0.05; table 1). In MCR2, which was located between MDR4 and MDR5, the mutation rate in nontumor tissues was also higher with boundary significance than in tumor tissues (0.0485 ± 0.1004 vs. 0.0121 ± 0.0485, p = 0.065). Throughout the core protein, only MCR6 and MDR8 showed slightly higher mutation numbers in the tumor group than in the nontumor group.
Discussion

The HBV core gene encodes a single polypeptide of 183 AAs (21 kDa) that assembles to form nucleocapsid particles. The core protein consists of two distinct domains: the N-terminal assembly domain (AA 1–149) that forms the contiguous spherical shell and the C-terminal protamine domain (AA 150–183) that is responsible for RNA packaging [19, 20]. As the C-terminal protamine domain is quite conserved as reported by the previous studies [10, 21, 22], we analyzed only the nucleotide substitutions from nt. 1901 to 2365 in the present investigation. This DNA fragment encodes the N-terminal 155 AAs of HBcAg, which contains the major known epitopes for CD4+ T cell, cytotoxic T lymphocyte (CTL) and B cell recognition [23–25]. While a large number of mutations in this region were observed in our study, they were mainly confined to three major and three minor MCRs. The average mutation rate in the MCRs was 0.0433 ± 0.0381, which was significantly higher than that in the other ‘MDRs’ (p < 0.001). These data indicate that the distribution of mutation in HBcAg is not random.

Most of the previous studies on core gene mutations were based on serum samples of chronic hepatitis patients [9–12, 21–23, 26, 27]; examples of molecular changes of HBcAg derived from the tumor tissues of HCC patients remain largely unexplored. To our knowledge, only Hosono et al. [28] has examined HBcAg sequences in 15 HCC tissues. They found that HBcAg mutations were predominantly within three major domains, i.e. codons 1–21, codons 80–121 and codons 130–135. Our results are compatible with that study. We noticed two hot-spot mutations (P5T/A/V and V13A/G/M/S) in codons 1–21 and two major MCRs (MCR3 and MCR4) in codons 80–121. The third major MCR, MCR5 in our study, resided precisely in codons 130–135. These data suggest that the mutation pattern of HBcAg isolated from HCC tissues may share some common features. Compared with the HBcAg sequences derived from the sera of non-HCC patients [9–12, 21–23, 26, 27, 29], those from HCC tissues displayed more AA changes at the N-terminus of HBcAg. Ohto’s group reported that in patients with severe liver damage (including chronic active hepatitis, cirrhosis, fulminant hepatitis and acute exacerbation to liver failure), HBcAg mutations were frequently found in codons 48–60 and codons 84–101 [10, 21, 22]. Hur et al. [12] later confirmed in a longitudinal observation that during the progression of chronic active hepatitis, the mutations gradually accumulated in a small segment of codons 84–97. Akarca and Lok [9] reported that the mutations associated with active liver disease were clustered in the middle of HBcAg, i.e. codons 59–66, codons 87–100, codons 125–135 and codons 147–155. Carman et al. [26] reported that the HBcAg mutations were concentrated in codons 74–84 in progressive liver disease and in codons 50–69 during clinical remission. In none of these studies was the N-terminus of HBcAg considered as a hypervariable region. However, in the present study, nine mutations were found in the first 38 AAs of HBcAg, particularly in MCR1 (codons 21–38). These N-terminal mutations either appear shortly before the occurrence of HCC or occur specifically in HCC cases and, therefore, could not be observed in patients at the chronic hepatitis stage. Further investigation with the use of sequential serum samples of HCC patients will be required to clarify the temporal relationship between the presence of HBcAg N-terminal mutation and HCC.

The HBV core protein is a principle target for immune response. It contains numerous epitopes for human CD4+ T cell (AA 1–20, AA 50–69 and AA 127–133), CTL (AA 18–27, AA 88–96 and AA 141–151) and B cell recognition (AA 74–89, AA 107–118 and AA 127–133) [30]. In chronic hepatitis patients, mutations have been found to preferentially locate within or flanking the dominant epitopes. In this study, HBcAg mutations were found to affect both T and B epitopes. Of the three major MCRs, MCR3 and MCR5 were in B cell epitopes (AA 74–89 and AA 127–133), while one (MCR4) overlapped with a CTL epitope (AA 88–96). Among the three minor MCRs, one (MCR2) was in a T helper epitope (AA 50–69) and two (MCR1 and MCR6) overlapped with CTL epitopes (AA 18–27 and AA 141–151, respectively). Each of the seven internal deletions affected a CTL epitope (AA 88–96) and a B cell epitope (AA 74–89) in the middle of HBcAg. As most of the HBcAg mutations occurred after the start of elevation of serum ALT levels and hepatitis B e antigen seroconversion [13, 23, 31], it was generally believed that the mutations were a result of selective pressure from the immune response. Whether HBcAg substitutions cause severe disease by inducing a naive immune response or merely immune escape has not been resolved. The missense mutations clustered in the CTL epitopes may break the immune tolerance and elicit stronger CTL-mediated immune clearance. Alternatively, mutations clustered in B-cell epitopes may result from an ineffective anti-HBc B cell response and may account for the ongoing disease and selection of mutations after seroconversion. In addition, mutations in the major T helper epitopes have been proposed to allow immune escape and minimize immune-mediated hepatitis, thus leading to clinical remis-
Very few studies have simultaneously examined the mutations of HBcAg in tumor and surrounding nontumor tissues. HCC children have been reported to have more mutations in the core gene than HBV chronic carriers [14]. However, in this study we found that the missense mutations in tumor tissues were significantly lower than those in adjacent nontumor tissues (p < 0.05), probably because the viruses in the tumor have had a reduced replication activity due to the unfavorable microenvironment. In contrast, those in the adjacent nontumor liver tissue retain a high replication rate and have had opportunity to accumulate more errors. Alternatively, in cases where replicating viruses were at a low level, the sequences isolated from HCC tissues may represent the integrated form of HBV DNAs. As HBV integration could occur at any stage of the infection [35, 36], it is reasonable to assume that the integrated HBV DNA contained fewer mutations compared with the replicative form of HBV in nontumor tissues. These hypotheses will need to be tested in the future by examining the HBV DNA level as well as the integration state of HBV in both tumor and adjacent nontumor liver tissues.

The limitation of this study is the lack of liver biopsies from patients at different stages of chronic hepatitis. No such sequence information in the previous literature was available to us to make the comparison. Therefore, the exact role of mutations, particularly those located in the N-terminus of HBcAg, in the development of HCC requires further study to clarify. Nevertheless, our data highlight the nature of the HBcAg mutation in tumor tissues of patients from an HCC-prevalent area. Occurrence of AA substitutions in MCRs might be used as an indicator of disease progression.

Acknowledgement

This work was supported by the Chinese State Key Project Specialized for Infectious Diseases (No. 2008ZX10002-015).

References

HBcAg Mutations in HCC Tissues

Intervirology 2012;55:29–35


