Int J Clin Exp Pathol 2017;10(12):11857-11866
www.ijcep.com /ISSN:1936-2625/IJCEP0068002

*Original Article*

# Associations between serum HBX quasispecies and their integration in hepatocellular carcinoma

Xiang Fang, Hang-Hang Wu, Jing-Jing Ren, Hai-Zhou Liu, Ke-Zhi Li, Ji-Lin Li, Yan-Ping Tang, Chan-Chan Xiao, Tian-Ren Huang, Wei Deng

*Department of Experimental Research, Affiliated Tumor Hospital of Guangxi Medical University, Nanning 530021, Guangxi Zhuang Autonomous Region, China*

Received October 26, 2017; Accepted November 10, 2017; Epub December 1, 2017; Published December 15, 2017

**Abstract:** HBV quasispecies are closely related to the course and outcome of liver disease. However, whether the complexity and diversity of HBX quasispecies affects its integration in the liver cell and thereby enhances the resultant carcinogenesis is still not clear. 15 HCC patients were recruited; genomic DNA and HBV DNA were extracted from liver cancer tissue and serum respectively. The integrated HBX fragment in liver cancer tissue was amplified by Alu repeat sequence-polymerase chain reaction (Alu-PCR) and sequenced. The serum HBX gene was amplified by nested PCR and sequenced. Quasispecies complexity and diversity, phylogenetic characteristics, lymphocyte count and survival time between HBX-integrated and HBX-unintegrated patients were evaluated. Results showed that the integrated HBX fragment was detected in the tumor tissue of nine patients, and the integration rate was 60.00% (9/15). Compared with the HBX-unintegrated patients, the HBX-integrated patients had a higher quasispecies complexity ($P$=0.028 and 0.004, at the nucleotide and amino acid levels, respectively). The HBX-integrated patients had a tendency of higher quasispecies diversity, lower lymphocyte count and the survival time. A total of 12 mutation sites were revealed in the HBX-integrated fragment after alignment with the reference sequence. In these, the HBX-integrated groups had significantly higher mutation frequencies at C1497T, A1630G, G1721A, A1762T/G1764A and A1774G. This study revealed influence factors of HBX integration both in virus and the host. The increased complexity and diversity of HBX quasispecies might destroy the host immune balance, and lead to HBX integration ultimately.

**Keywords:** Hepatitis B virus X gene, hepatocellular carcinoma, integration, quasispecies

## Introduction

Hepatocellular carcinoma (HCC) is the fifth most common malignancy that causes tumor-related deaths worldwide. Approximately 600,000 people die from HCC every year, and nearly half of them occurred in China [1]. Chronic hepatitis B virus (HBV) infection is the main cause of HCC in China. HBV is a double-stranded circular DNA virus. During HBV replication, the reverse transcriptase lacks the 3'-5' correction function, and these results in a high mismatch rate in base-pairs. Hence, a virus variant is one clone of the HBV, and a large number of different variants are generated as HBV quasispecies under natural selection. The concept of HBV quasispecies was first proposed in 1993 [2]. The complexity and diversity of HBV quasispecies are reliable parameters for managing HBV infection and HCC. Many studies have confirmed that the characteristics of HBV quasispecies are closely related to the course and outcome of liver disease [3, 4].

The HBV gene is 3200 bp long. The genome contains four open reading frames (ORFs), namely, S, C, P and X, of which the HBV X gene (HBX) is the smallest. The X gene is 465 bp long, and its coding region is between 1374 and 1838 nucleotides. The X protein encoded by the HBX is composed of 154 amino acids. HBX is the most frequently integrated fragment in the HBV genome. In addition, HBX and its product play key roles in the occurrence and development of hepatitis and HCC by its integration in the genomic DNA of liver cells [5-7].

However, whether the complexity and diversity of HBX quasispecies affects its integration in

**Table 1.** The primer sequences of Alu-PCR

| Primer name | Primer sequence | Note |
|---|---|---|
| UP5 | 5'CAGUGCCAAGUGUUUGCUGACGCCAAAGUGCUGGGAUUA3' | Alu-forward sequence |
| T3-515 | 5'AUUAACCCUCACUAAAGCCUCGAUAGAUYRYRCCAYUGCAC3' | Alu-reverse sequence |
| UP6 | 5'CAAGTGTTTGCTGACGCCAAAG3' | Alu-forward sequence |
| MidT3 | 5'ATTAACCCTCACTAAAGCCTCG3' | Alu-reverse sequence |
| pUTP | 5'ACAUGAACCUUUACCCCGUUGC3' | HBX1 |
| MM37 | 5'TGCCAAGTGTTTGCTGACGC3' | HBX2 |
| MM60 | 5'CTGCCGATCCATACTGCGGAAC3' | HBX3 |

Note: HBX1: nucleotides 1131-1152nt; HBX2: nucleotides 1174-1193nt; HBX3: nucleotides 1258-1279nt.

the liver cell and thereby enhances the resultant carcinogenesis is still not clear [8, 9]. In the present study, the HBX was used as the target region to explore the associations between serum HBX quasispecies and their integration, and ultimately reveal the underlying carcinomatosis mechanism of HBX.

## Materials and methods

### *Patient characteristics*

15 HCC patientsin Barcelona Clinic Liver Cancer (BCLC) stage A1 who underwent HCC resection at the Affiliated Tumor Hospital of Guangxi Medical University between 2010 and 2016 were selected for this study. All patients had hepatitis B surface antigen positive for more than 6 months, and tested negative for HCV, HIV and HDV. The diagnoses of HCC were based on the pathological finding of HCC cells. Liver tissue from the HCC lesions and peripheral whole blood samples were collected respectively after informed consent of all subjects. The clinical data collected were HBeAg status, HBsAb status, TbiL (total bilirubin), ALT (alanine aminotransferase), AST (aspartate aminotransferase) AFP (alpha fetoprotein), serum HBV DNA and lymphocyte count ($CD3^+$ T lymphocyte, $CD4^+$ T lymphocyte, $CD8^+$ T lymphocyte, CD4/CD8, $CD3^-CD16^+56^+$ natural killer cells and $CD19^+$ B lymphocyte). All patients were followed-up after surgery until June 30, 2017. The study was approved by the Ethics Committee of the Affiliated Tumor Hospital of Guangxi Medical University.

### *Sample preparation*

All tissue and blood samples were stored at -80°C until use. Genomic DNA from 0.2 g of liver tissue was extracted using DN easy Blood & Tissue Kit (QIAGEN) for Alu-PCR. The peripheral whole blood was centrifuged to separate serum. HBV DNA was extracted from 200 μL serum using QIA amp MinElute Virus Spin Kit (QIAGEN) for nested-PCR. All the experimental operations were carried out according to the manufacturer's instructions.

### *Detection integrated-HBX (Int-HBX) by Alu-PCR*

According to the HBX sequence and human Alu repeat, special primers were designed and are shown in **Table 1** for Alu-PCR, to amplify the integrated fragments combined target HBX and the adjacent cell gene sequences using the genomic DNA as template. The first set of PCR amplification primers included UP5/T3-515. The PCR system comprised a reaction volume of 100 μL (50 μL of Taq PCR Master Mix, 2 μL of 10 mmol/L pUTP mix, 2 μL of primer UP5/T3-515, 3 μL of HBV DNA and 43 μL of deionized water). The PCR procedure was as follows: pre-denaturation at 94°C for 3 min followed by 35 cycles of denaturation at 94°C for 30 s, annealing at 56°C for 30 s and extension at 72°C for 1 min and a final extension at 72°C for 10 min. The product was digested by 5 U UDG enzyme at 37°C for 60 min to damage the DNA site containing dUTP. Then, the digestion was terminated and the reaction products were incubated at 95°C for 10 min. The second set of PCR amplification primers included MM37 and UP6/MidT3. The PCR system comprised a reaction volume of 100 μL (50 μL of Taq PCR Master Mix, 2 μL of primer MM37, 2 μL of primer UP6/MidT3, 3 μL of HBV DNA and 43 μL of deionized water). The PCR procedure was as follows: pre-denaturation at 94°C for 3 min followed by 35 cycles of denaturation at 94°C for 30 s, annealing at 48°C for 30 s and extension at 72°C for 1 min and a fi-

**Table 2.** Clinical data of the patients in this study

| Characteristic | HBX-integration (n=9) | HBX-unintegration (n=6) | *P* |
|---|---|---|---|
| Age (years) | 44.22±10.45 | 45.50±13.85 | 0.841 |
| Gender (males/females) | 7/2 | 5/1 | 1.000 |
| HBeAg (+) | 8 | 6 | 1.000 |
| HBsAb (+) | 7 | 5 | 1.000 |
| ALT (U/L) | 54.89±18.47 | 61.83±10.18 | 0.767 |
| AST (U/L) | 69.11±13.06 | 68.50±18.39 | 0.985 |
| TbiL (μmol/L) | 13.00±5.43 | 12.05±4.04 | 0.721 |
| AFP (ng/ml) | 654.13±46.82 | 559.30±48.41 | 0.172 |
| HBV DNA (log IU/mL) | 3.88±1.91 | 4.43±0.89 | 0.527 |

Note: n: The total number of patients; ALT: alanine aminotransferase; AST: aspartate aminotransferase; TBniL: Total bilirubin; AFP: alpha fetoprotein.

**Table 3.** The integration of HBX in cancer tissue

| Patient | Integrated chromosome | Integrated gene |
|---|---|---|
| 1 | 1p36.32 | P73 |
| 2 | 12P13.33 | WNT5B |
| 3 | 17p13.1 | P53 |
| 4 | 17p13.1 | TP53 |
| 5 | 19q13.32 | BSG |
| 6 | 5p15.33 | hTERT |
| 7 | 2p24.2 | GEN1 |
| 8 | 5q13.2 | Cyclin B1 |
| 9 | 1p36.12 | E2F2 |

nal extension at 72°C for 10 min. The amplicons were diluted 500 times as the third round of PCR template. The third set of "half nest" PCR amplification primers included MM60 and UP6/MidT3. The PCR system comprised a reaction volume of 100 μL (50 μL of Taq PCR Master Mix, 2 μL of primer MM60, 2 μL of primer UP6/MidT3 and 46 μL of DNA template). The PCR procedure was as follows: pre-denaturation at 94°C for 3 min followed by 35 cycles of denaturation at 94°C for 30 s, annealing at 60°C (forward)/56°C (reverse) for 30 s and extension at 72°C for 1 min and a final extension at 72°C for 10 min. After that, the end-products of Alu-PCR were sent directly to the Beijing Genomics Institute (BGI) for sequencing. The sequencing results were applied to determine the positions where HBX integrated in the human genome and the corresponding target genes by using NCBI BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi). The HCC patients were divided into two groups (the HBX-integrated group and the HBX-unintegrated group) according to whether HBX integration in liver cells was observed.

*Amplification of HBX by nested PCR*

Using HBV DNA extracted from serum as template, HBX was amplified by nested PCR. The first set of PCR primers (outer primer pair) included f2 (5'-CAAGTGTTTGCTACGCAACC-3') and r2 (5'-ACAGCTTGGAGGCTTGAACAG-3'). The first round of PCR contained a reaction volume of 100 μL (10 μL of 10× buffer, 2 μL of 10 mmol/L dNTP mix, 5 μL of primer f2, 5 μL of r2, 0.5 μL of tap enzyme, 10 μL of HBV DNA and 67.5 μL of deionized water. The PCR procedure was as follows: pre-denaturation at 95°C for 5 min followed by 40 cycles of denaturation at 94°C for 1 min, annealing at 44.5°C for 1 min and extension at 72°C for 1 min and a final extension at 72°C extension for 10 min. The second set of PCR primers (inner primer pair) included f1a (5'-TCCTTCCCATGGCTGCTCGGGTGTGCTG-3') and r1a (5'-CATGAGATGATTAGGCAGAGGTGAAAAAG-3'). The first round of PCR involved a reaction volume of 100 μL (10 μL of 10× buffer, 2 μL of 10 mmol/L dNTP mix, 5 μL of primer f1a, 5 μL of r1a, 0.5 μL of tap enzyme, 4 μL of HBV DNA and 73.5 μL of deionized water). The PCR procedure was as follows: pre-denaturation at 95°C for 5 min followed by 35 cycles of denaturation at 94°C for 1 min, annealing at 53°C for 1 min and extension at 72°C for 1 min and a final extension at 72°C for 10 min. The end-product was 465 bp long and covered the coding region of HBX between nucleotides 1374 and 1838.

*Purification and cloning*

The nested PCR products were separated by 1% agarose gel electrophoresis. The target bands were excised under a long-wave UV lamp and purified by Midi Purification Kit (TIANGEN, China) according to the manufacturer's instructions. The purified amplicons were cloned into PUCM-T vectors and transformed into Escherichia coli DH5α competent cells using

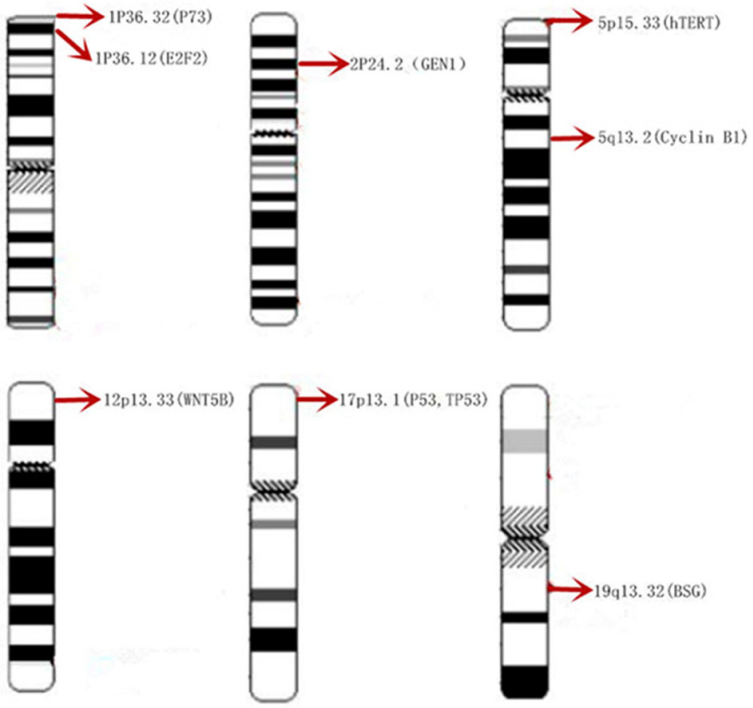


**Figure 1.** The distribution of HBX integration sites in human chromosomes.

the T vector PCR product cloning kit (Sangon, China) according to the manufacturer's instructions. A total of 15-25 positive clones of each sample were selected for growth in LB liquid medium (24 h at 37°C, 200 rmp), and the bacterial pellets were sent to the Beijing Genomics Institute (BGI) for sequencing, to obtained the HBX quasispecies sequences.

*HBX quasispecies analysis*

All cloned nucleotide sequences of HBX were input into MEGA 6.0 software, and a phylogenetic tree was constructed for observing visually the genetic relationships between the different HBX variants, namely HBX quasispecies.

In addition, the heterogeneity of the quasispecies, including complexity and diversity, was evaluated at the nucleotide and amino acid levels. Quasispecies complexity was expressed as the Shannon entropy (Sn) [10], which refers to the proportion of different virus quasispecies. Sn ranges from 0 to 1; 0 means that all sequences are identical and 1 means that each sequence is different and unique. Sn was calculated as $Sn=-\sum i\ (pi\ lnpi)/lnN$, where pi is the frequency of each sequence in the virus population and N is the total clone number. The diversity of quasispecies was described by the average genetic distance (d, also known as the Hamming distance), the number of synonymous substitutions per site (dS) and the number of non-synonymous substitutions per site (dN). The diversity of quasispecies was calculated by MEGA 6.0 software, where d was calculated under the Kimura 2-Parameter model and dn and ds were calculated under the Jukes-Cantor model [11].

Finally, the HBX-integrated fragments in liver cancer tissues were aligned with the HBV reference sequence (Accession Version NC-003977.1 in NCBI) in order to find out which of the mutation sites might affect integration. Then the distribution of these mutations sites in serum HBX quasispecies were analyzed using the input data in the MEGA 6.0 software. The mutation frequencies between the HBX-integration group and the HBX-unintegration group were then compared to analyze the relationship of the HBX mutation in serum and the integration in liver tissues.

*Statistical analysis*

SPSS 19.0 software was used for statistical analysis. The mean comparison was performed using Student's t-test. Rate comparison was performed using the chi-square test. Survival analyses were performed using the Kaplan-Meier method and Log Rank test. Statistical significance was considered at $P<0.05$.

## Results

*HBX integration rate and comparisons of clinical characteristics*

Of all the 15 subjects, 9 patients (cases number 1 to 9) were detected with an integrated fragment of the HBX gene in the lesion of HCC, with an integrated rate of 60.00% (9/15). The comparisons of clinical characteristics between

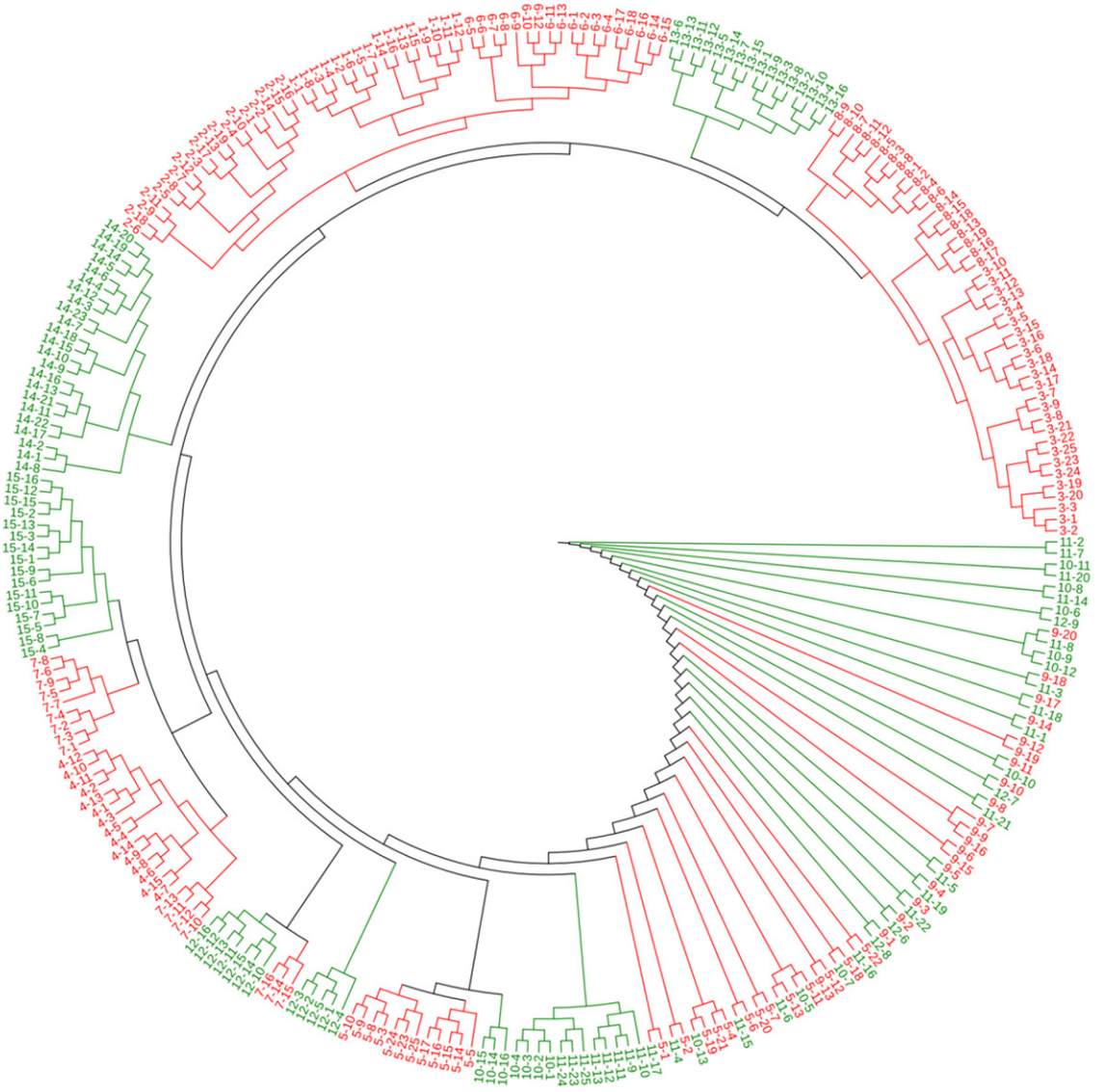

**Figure 2.** Phylogenetic trees were constructed for clone sequences with a Bootstrap method. "-" The former number represents the case number, "-" after the number represents the clone number. A red branch represents the HBX-integrated group and a green branch represents the HBX-unintegrated group of patients.

HBX-integrated group and the HBX-unintegrated group are showed in **Table 2**. There was no significant difference in gender, age, HBeAg seropositive rate, HbsAb seropositive rate, ALT level, AST level, TbiL level, AFP level, and the HBV DNA level between two groups ($P$>0.05 for all). This suggested that the baseline characteristics between the two groups of patients were similar.

*The target gene of HBX integration in human genome*

**Table 3** shows the integration sites. The genes involved in the integration were P73, WNT5B, P53, TP53, BSG, hTERT, GEN1, Cyclin B1 and E2F2 (**Figure 1**).

*Phylogenetic analysis of HBX quasispecies*

A phylogenetic tree constructed with a total of 285 HBX quasispecies clone sequence are shown in **Figure 2**. The HBX-integrated group included 173 clones and the HBX-unintegrated group included 112. The clone sequences from the same patient were gathered, which shows the homologous nature of the HBX quasispecies. The branches of trees from the HBX-integrated group were generally longer than

**Table 4.** The comparison of quasispecies complexity and diversity between HBX-integrated and HBX-unintegrated groups of patients

| | HBX-integrated (N=137) | HBX-unintegrated Patients (N=112) | *P* value |
|---|---|---|---|
| SN (nt) | 0.590±0.145 | 0.437±0.049 | 0.028 |
| SN (aa) | 0.458±0.125 | 0.262±0.081 | 0.004 |
| d | 0.00556±0.00646 | 0.00300±0.00352 | 0.396 |
| dS | 0.00822±0.00687 | 0.00417±0.00691 | 0.284 |
| dN | 0.00456±0.00659 | 0.00233±0.00301 | 0.457 |

Note: SN (nt): Quasispecies complexity at the nucleotide level; SN (aa): Quasispecies complexity at the Amino acid level; d: Genetic distance; dS: the number of synonymous substitutions per synonymous site; dN: the number of non-synonymous substitutions per non-synonymous site.

**Table 5.** Comparison of the mutation frequencies of the HBX integration related mutations between the HBX-Integration group and the HBX-unintegration group of patients

| Mutation sites | HBX-Integration (N=173) | HBX-unintegration (N=112) | $X^2$ | *P* |
|---|---|---|---|---|
| T1425C | 61 (35.3%) | 39 (34.8%) | 0.006 | 0.940 |
| C1473T | 46 (26.6%) | 33 (28.6%) | 0.280 | 0.596 |
| C1484A | 154 (89.0%) | 105 (93.8%) | 1.837 | 0.175 |
| T1485C | 134 (77.5%) | 91 (81.3%) | 0.589 | 0.443 |
| C1497T | 79 (45.7%) | 39 (34.8%) | 5.593 | 0.018 |
| A1512G | 20 (11.6%) | 17 (15.2%) | 0.788 | 0.375 |
| A1544T | 72 (41.6%) | 41 (36.6%) | 0.714 | 0.398 |
| A1630G | 95 (54.9%) | 23 (20.5%) | 33.118 | 0.000 |
| G1635A | 61 (35.3%) | 38 (33.9%) | 0.053 | 0.818 |
| G1721A | 60 (34.7%) | 15 (13.4%) | 15.891 | 0.000 |
| A1762T/G1764A | 124 (71.7%) | 39 (34.8%) | 37.718 | 0.000 |
| A1774G | 53 (30.6%) | 19 (16.9%) | 6.730 | 0.009 |

Note: N: Number of clones.

those of trees from the HBX-unintegrated group.

*HBX quasispecies heterogeneity*

As shown in **Table 4**, HBX quasispecies complexity was significantly higher in the HBX-integrated groups (173 clones) when compared to the unintegrated groups (112 clones) at both the nucleotide level (Sn: 0.590±0.145 vs. 0.437±0.049, respectively; *P*=0.028) and at the amino acid level (Sn: 0.458±0.125 vs. 0.262±0.081, respectively; *P*=0.004). As for quasispecies diversity, d (0.00556±0.00646 vs. 0.00300±0.00352, respectively; *P*=0.396), dN (0.00456±0.00659 vs. 0.00233±0.00301, respectively; *P*=0.457), and dS (0.00822±0.00687 vs. 0.00417±0.00691, respectively; *P*=0.284) showed no significant differences between the two groups.

*Mutation site of HBX-integrated fragment and its distribution in HBX quasispecies*

A total 12 mutation sites were revealed in HBX-integrated fragment after alignment with the reference sequence, including synonymous mutations like T1425C, C1484A, A1544T, G1721A and A1774G, and non-synonymous mutations like C1473T (leucine mutations to phenylalanine), T1485C (serine mutations to proline), C1497T (proline mutations to serine), A1512G (threonine mutation to alanine), A1630G (histidine mutation to arginine), A1635G (isoleucine mutation to valine), A1762T/G1764A (lysine mutation to methionine and valine mutation to isoleucine ). The distribution of these mutation sites between the HBX-integrated groups (173 clones) and the unintegrated groups (112 clones) were compared. It was found that the HBX-integrated groups had significantly higher mutation frequencies at C1497T, A1630G, G1721A, A1762T/G1764A and A1774G. The other mutations showed no significant difference between the two groups (**Table 5**).

*HBX integration and cellular immunity*

Lymphocyte counts between the HBX-integrated group and the HBX-unintegrated group were compared (**Table 6**). Although the results were not statistically different, there was a tendency for the proportions of total T lymphocyte ($CD3^+$) and helper T-cells ($CD3^+CD4^+$), CD4/CD8 to be higher in the HBX-integrated group. The proportion of cytotoxic T-cells ($CD3^+CD8^+$) was higher in the HBX-unintegrated group and the natural killer cells ($CD3^-CD16^+56^+$) and B lymphocytes ($CD19^+$) were distributed in a similar way.

*HBX integration and survival time*

Survival analyses showed that mean for survival time of the HBX-integrated group and the HBX-unintegrated groups were 33.0 and 44.7 months respectively. **Figure 3** revealed the

**Table 6.** The comparison of lymphocyte counts between the HBX-integrated and HBX-unintegrated patients

| | $CD3^+$ (%) | $CD3^+CD4^+$ (%) | $CD3^+CD8^+$ (%) | $CD4^+/CD8^+$ (%) | $CD3^-CD16^+CD56^+$ (%) | $CD19^+$ (%) |
|---|---|---|---|---|---|---|
| HBX-integrated | 63.686±10.9271 | 36.643±7.0639 | 21.0571±5.2073 | 1.886±0.7381 | 19.971±11.6971 | 10.500±5.1085 |
| HBX-unintegrated | 69.100±9.1545 | 44.200±12.2079 | 18.940±5.3266 | 2.560±1.0621 | 19.220±7.8846 | 10.900±4.2544 |
| *P* | 0.388 | 0.203 | 0.507 | 0.221 | 0.904 | 0.889 |

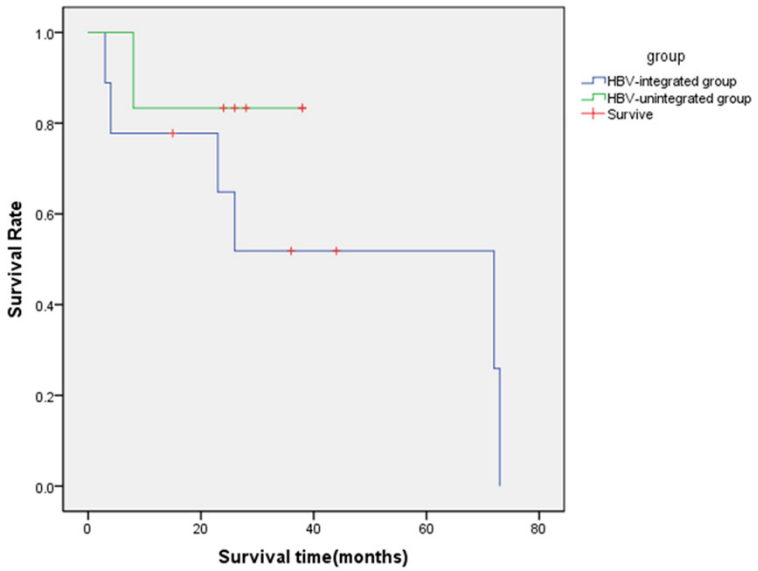


**Figure 3.** Survival curves of the HBX-integrated and HBX-unintegrated group of patients using the Kaplan-Meier method.

overall comparisons of survival time between two groups. It appears that the HBX-unintegrated patients lived longer. However, the Log Rank test showed this was not statistically significant (chi-square value was 1.151, *P*=0.283).

## Discussion

It is possible to detect HBV integration in 80%-90% of liver tissue in HBV related HCC patients [12, 13] and 57% of HBV integration fragment was shown to be HBX [14]. HBX is considered to integrate into the human chromosome and promote the occurrence and development of HCC, by regulating the gene function surrounding the integration site, expressing the carcinogenic HBX truncated protein and increasing the instability of human genes [15-17]. In several previous studies, the integration sites of HBV were found to be randomly distributed in human chromosomes [18, 19]. Nowadays, with the development of sequencing technology, more and more new integration sites have been discovered to be regularly distributed in human chromosomes.

Yang et al. found that in HCC tissues, HBV preferentially integrated in chromosomes 5, 8, 10 and 19 [20]. Using the massive anchored parallel sequencing (MPAS) technique, Dong et al. found that HBV was preferentially integrated into chromosome 17 and tended to be integrated into transcripts of the human genome [21]. The target genes adjacent to the HBX integration sites included TERT, MLL4, FN1, cyclin A, cyclin E1 and CCNE1 and these genes are involved in cell growth, proliferation, differentiation and immortalization [22, 23]. However, the viral and host factors that might affect HBV integration in the liver cell and the prognosis after HBV integration in HCC patients are still not clear.

In this study, we found that the integration rate of HBX in liver cancer tissue was 60% and the target genes involved in HBX integration were P73, WNT5B, P53, TP53, BSG, hTERT, GEN1, Cyclin B1 and E2F2. These genes are involved in cells growth, reproduction and some essential signal transduction pathways that are as-sociated with tumor occurrence, invasion and metastasis. For example, P53 and TP53 may be involved in the occurrence of HCC by regulating the transcription of sodium /iodide symporter [24]. The WNT signaling pathway is involved in the development of HCC by regulating cell growth, apoptosis, tissue remodeling and angiogenesis [25, 26]. HBX integration into human telomerase reverse transcriptase (TERT) gene can increase TERT

expression [27], and TERT is a gene that can control telomerase expression. TERT mutations have been shown to be a common event for many tumor types (including HCC), and may be a biomarker for the early detection of HCC [28].

When HCC patients were followed up in this study, we found that the HBX-unintegrated group had a longer survival time than HBX-integrated group although this was not statistically significant. This result suggested that HBX integration might influence the prognosis of HCC patients although the underlying mechanisms of this relationship remain unclear. We speculate that the HBX integration may cause the pathogenesis of HCC by regulating the function of target genes, which are associated with the growth, signal transduction, proliferation and apoptosis of hepatocytes.

Since the integration of HBX is closely related to the occurrence of HCC, it is worth exploring the factors that influence the integration of HBX from the perspective of both the virus and the host. Shannon entropy is an important parameter used to express the complexity of HBV quasispecies, which has been widely used to study HBV-related liver disease [11, 29].

In this study, the HBX-integrated patients had a significantly higher quasispecies complexity than the HBX-unintegrated patients, both at the nucleotide and amino acid levels. For quasispecies diversity, the HBX-integrated group had a higher trend than the HBX-unintegrated group, and the phylogenetic tree analysis showed that the red branch (representing the HBX-integrated group) was generally longer than the green branch (representing the HBX-unintegrated group). This indicates a higher gene diversity in the HBX-integrated group. The increased HBX quasispecies complexity and diversity may cultivate a larger genetic pool, resulting in a higher chance to form the possible predominant quasispecies and integrated in human genome. With regard to the host, increased quasispecies complexity and diversity may trigger an enhanced immune-mediated host response [30] in which T cells might play a key role, as demonstrated by recent studies that tumor-infiltrating CD4 T cells are impaired in patients with HCC [31]. Our study found that the total T lymphocyte ($CD3^+$) and helper T-cells ($CD3^+CD4^+$) in the HBX-unintegrated group of patients tended to be lower than the HBX-unintegrated group. This result suggested that the increased complexity and diversity of HBX quasispecies, to some extent, may destroy the host immune balance, and the virus-host interplay and this may lead to HBX integration. More *in vitroand in vivo* studies are needed to prove this hypothesis.

When focused on the mutation of HBX quasispecies, we found that the HBX-integrated group had significantly higher mutation frequencies at C1497T, A1630G, G1721A, A1762T/G1764A and A1774G. A1630G and G1721A were located in the enhancer of the HBV gene. To our knowledge, two enhancers (Enh I and Enh II) have been found in the HBV gene. Enhancers play a regulatory role in the transcription and replication of HBV. Enh II consists of two fragments, namely, the A (nt l627-nt1686) and B fragments (nt1687-nt1774). Previous studies have confirmed that the nt-1632-nt1724 nucleotide mutation is associated with the occurrence of liver cancer. G1721A, A1762T/G1764A and A1774G are located in the BCP/precore regions of the HBV genome and BCP mutations have a significant association with HCC [32]. The A1762T/G1764A mutations are important risk factors for HCC [33]. This study indicated that C1497T, A1630G, G1721A, A1762T/G1764A and A1774G may be important factors for the formation of predominant quasispecies associated with HBV integration. These predominant quasispecies are more likely to integrate with human chromosomes to promote the occurrence of HCC.

To the best of our knowledge, this study is the first to elucidate the associations between serum HBX quasispecies and their corresponding liver cancer tissue HBX integration, which is of great significance in order to reveal the influential factors associated with HBX integration both in the virus and the host. For HCC patients infected with HBV, clinical attention should be paid to monitor their immune function and HBV characteristics. In addition, our data would suggest that some measures should be taken to improve immunity and reduce the HBV quasispecies complexity and diversity, especially for those patients infected with predominant HBX quasispecies containing carcinogenic mutations. This may help to

reduce the incidence of HBV integration and extend the life expectancy of patients with HCC.

## Acknowledgements

This study was supported by grants from the National Natural Science Foundation of China (Nos. 81660561 and 81260319). The authors declare that they have no competing interests. The authors would like to thank Dr. Dev Sooranna, Imperial College London, for editing the manuscript.

## Disclosure of conflict of interest

None.

**Address correspondence to:** Drs. Wei Deng and Tian-Ren Huang, Department of Experimental Research, Affiliated Tumor Hospital of Guangxi Medical University, 22 Shuangyong Road, Nanning 530021, Guangxi Zhuang Autonomous Region, China. Tel: 86-771-5322502; 86-13481085969; Fax: 86-771-5322502; E-mail: dengwei@gxmu.edu.cn (WD); Tel: 86-771-5322502; 86-13607862198; Fax: 86-771-5322502; E-mail: tianrenhuang@sina.com (TRH)

## References

[1] Siegel RL, Miller KD and Jemal A. Cancer statistics, 2016. CA Cancer J Clin 2016; 66: 7-30.

[2] Carman W, Thomas H and Domingo E. Viral genetic variation: hepatitis B virus as a clinical example. Lancet 1993; 341: 349-53.

[3] Banerjee P, Chakraborty A, Mondal RK, Khatun M, Datta S, Das K, Pandit P, Mukherjee S, Banerjee S, Ghosh S, Chakrabarti S, Chowdhury A and Datta S. HBV quasispecies composition in Lamivudine-failed chronic hepatitis B patients and its influence on virological response to Tenofovir-based rescue therapy. Sci Rep 2017; 7: 44742.

[4] Xue Y, Wang MJ, Huang SY, Yang ZT, Yu DM, Han Y, Zhu MY, Huang D, Zhang DH, Gong QM and Zhang XX. Characteristics of CpG Islands and their quasispecies of full-length hepatitis B virus genomes from patients at different phases of infection. Springerplus 2016; 5: 1630.

[5] Wang M, Xi D and Ning Q. Virus-induced hepatocellular carcinoma with special emphasis on HBV. Hepatol Int 2017; 11: 171-80.

[6] Ching RHH, Sze KMF, Lau EYT, Chiu YT, Lee JMF, Ng IOL and Lee TKW. C-terminal truncated hepatitis B virus X protein regulates tumorigenicity, self-renewal and drug resistance via STAT3/Nanog signaling pathway. Oncotarget 2017; 8: 23507-23516.

[7] Levrero M and Zucman-Rossi J. Mechanisms of HBV-induced hepatocellular carcinoma. J Hepatol 2016; 64 Suppl: S84-S101.

[8] Lee S, Lee MJ, Zhang J, Yu GR and Kim DG. C-terminal-truncated HBV X promotes hepatoncogenesis through inhibition of tumor-suppressive beta-catenin/BAMBI signaling. Exp Mol Med 2016; 48: e275.

[9] Takeda H, Takai A, Inuzuka T and Marusawa H. Genetic basis of hepatitis virus-associated hepatocellular carcinoma: linkage between infection, inflammation, and tumorigenesis. J Gastroenterol 2017; 52: 26-38.

[10] Fishman SL and Branch AD. The quasispecies nature and biological implications of the hepatitis C virus. Infect Genet Evol 2009; 9: 1158-67.

[11] Zhang AY, Lai CL, Poon RT, Huang FY, Seto WK, Fung J, Wong DK and Yuen MF. Hepatitis B virus full-length genomic mutations and quasispecies in hepatocellular carcinoma. J Gastroenterol Hepatol 2016; 31: 1638-45.

[12] Hai H, Tamori A and Kawada N. Role of hepatitis B virus DNA integration in human hepatocarcinogenesis. World J Gastroenterol 2014; 20: 6236-43.

[13] Sung WK, Zheng H, Li S, Chen R, Liu X, Li Y, Lee NP, Lee WH, Ariyaratne PN, Tennakoon C, Mulawadi FH, Wong KF, Liu AM, Poon RT, Fan ST, Chan KL, Gong Z, Hu Y, Lin Z, Wang G, Zhang Q, Barber TD, Chou WC, Aggarwal A, Hao K, Zhou W, Zhang C, Hardwick J, Buser C, Xu J, Kan Z, Dai H, Mao M, Reinhard C, Wang J and Luk JM. Genome-wide survey of recurrent HBV integration in hepatocellular carcinoma. Nat Genet 2012; 44: 765-69.

[14] Jiang S, Yang Z, Li W, Li X, Wang Y, Zhang J, Xu C, Chen PJ, Hou J, McCrae MA, Chen X, Zhuang H and Lu F. Re-evaluation of the carcinogenic significance of hepatitis B virus integration in hepatocarcinogenesis. PLoS One 2012; 7: e40363.

[15] Cao S, Wendl MC, Wyczalkowski MA, Wylie K, Ye K, Jayasinghe R, Xie M, Wu S, Niu B, Grubb R 3rd, Johnson KJ, Gay H, Chen K, Rader JS, Dipersio JF, Chen F and Ding L. Divergent viral presentation among human tumors and adjacent normal tissues. Sci Rep 2016; 6: 28294.

[16] Ng KY, Chai S, Tong M, Guan XY, Lin CH, Ching YP, Xie D, Cheng AS and Ma S. C-terminal truncated hepatitis B virus X protein promotes hepatocellular carcinogenesis through induction of cancer and stem cell-like properties. Oncotarget 2016; 7: 24005-17.

[17] Buendia MA and Neuveut C. Hepatocellular carcinoma. Cold Spring Harb Perspect Med 2015; 5: a021444.

[18] Matsubara K and Tokino T. Integration of hepatitis B virus DNA and its implications for hepatocarcinogenesis. Mol Biol Med 1990; 7: 243-60.

[19] Jiang Z, Jhunjhunwala S, Liu J, Haverty PM, Kennemer MI, Guan Y, Lee W, Carnevali P, Stinson J, Johnson S, Diao J, Yeung S, Jubb A, Ye W, Wu TD, Kapadia SB, de Sauvage FJ, Gentleman RC, Stern HM, Seshagiri S, Pant KP, Modrusan Z, Ballinger DG and Zhang Z. The effects of hepatitis B virus integration into the genomes of hepatocellular carcinoma patients. Genome Res 2012; 22: 593-601.

[20] Yang X, Wu L, Lin J, Wang A, Wan X, Wu Y, Robson SC, Sang X and Zhao H. Distinct hepatitis B virus integration patterns in hepatocellular carcinoma and adjacent normal liver tissue. Int J Cancer 2017; 140: 1324-30.

[21] Ding D, Lou X, Hua D, Yu W, Li L, Wang J, Gao F, Zhao N, Ren G, Li L and Lin B. Recurrent targeted genes of hepatitis B virus in the liver cancer genomes identified by a next-generation sequencing-based approach. PLoS Genet 2012; 8: e1003065.

[22] Homs M, Caballero A, Gregori J, Tabernero D, Quer J, Nieto L, Esteban R, Buti M and Rodriguez-Frias F. Clinical application of estimating hepatitis B virus quasispecies complexity by massive sequencing: correlation between natural evolution and on-treatment evolution. PLoS One 2014; 9: e112306.

[23] Li X, Zhang J, Yang Z, Kang J, Jiang S, Zhang T, Chen T, Li M, Lv Q, Chen X, McCrae MA, Zhuang H and Lu F. The function of targeted host genes determines the oncogenicity of HBV integration in hepatocellular carcinoma. J Hepatol 2014; 60: 975-84.

[24] Guerrieri F, Piconese S, Lacoste C, Schinzari V, Testoni B, Valogne Y, Gerbal-Chaloin S, Samuel D, Brechot C, Faivre J and Levrero M. The sodium/iodide symporter NIS is a transcriptional target of the p53-family members in liver cancer cells. Cell Death Dis 2013; 4: e807.

[25] Huang S, Xie Y, Yang P, Chen P and Zhang L. HCV core protein-induced down-regulation of microRNA-152 promoted aberrant proliferation by regulating Wnt1 in HepG2 cells. PLoS One 2014; 9: e81730.

[26] Tomimaru Y, Xu CQ, Nambotin SB, Yan T, Wands JR and Kim M. Loss of exon 4 in a human T-cell factor-4 isoform promotes hepatic tumourigenicity. Liver Int 2013; 33: 1536-48.

[27] Toh ST, Jin Y, Liu L, Wang J, Babrzadeh F, Gharizadeh B, Ronaghi M, Toh HC, Chow PK, Chung AY, Ooi LL and Lee CG. Deep sequencing of the hepatitis B virus in hepatocellular carcinoma patients reveals enriched integration events, structural alterations and sequence variations. Carcinogenesis 2013; 34: 787-98.

[28] Pezzuto F, Buonaguro L, Buonaguro FM and Tornesello ML. Frequency and geographic distribution of TERT promoter mutations in primary hepatocellular carcinoma. Infect Agent Cancer 2017; 12: 27.

[29] Zhou TC, Li X, Li L, Li XF, Zhang L and Wei J. Evolution of full-length genomes of HBV quasispecies in sera of patients with a coexistence of HBsAg and anti-HBs antibodies. Sci Rep 2017; 7: 661.

[30] Lim SG, Cheng Y, Guindon S, Seet BL, Lee LY, Hu P, Wasser S, Peter FJ, Tan T, Goode M and Rodrigo AG. Viral quasi-species evolution during hepatitis be antigen seroconversion. Gastroenterology 2007; 133: 951-58.

[31] Zhou G, Sprengers D, Boor PPC, Doukas M, Schutz H, Mancham S, Pedroza-Gonzalez A, Polak WG, de Jonge J, Gaspersz M, Dong H, Thielemans K, Pan Q, IJzermans JNM, Bruno MJ and Kwekkeboom J. Antibodies against immune checkpoint molecules restore functions of tumor-infiltrating T cells in hepatocellular carcinomas. Gastroenterology 2017; 153: 1107-19.

[32] Mendy ME, Kaye S, Le Roux E, Kirk GD, Jeng-Barry A, McConkey S, Cotten M, Kuniholm MH, Leligdowicz A, Hainaut P, Rowland-Jones S and Whittle H. Application of a novel, rapid, and sensitive oligonucleotide ligation assay for detection of cancer-predicting mutations in the precore and basal core promoter of hepatitis B virus. J Clin Microbiol 2008; 46: 2723-30.

[33] Yang Z, Zhuang L, Lu Y, Xu Q, Tang B and Chen X. Naturally occurring basal core promoter A1762T/G1764A dual mutations increase the risk of HBV-related hepatocellular carcinoma: a meta-analysis. Oncotarget 2016; 7: 12525-12536.