Quantitative evaluation of hepatitis B virus mutations and hepatocellular carcinoma risk: a meta-analysis of prospective studies

Yang Yang¹, Jiang-Wei Sun², Long-Gang Zhao², Freddie Bray³, Yong-Bing Xiang²

¹School of Public Health, Fudan University, Shanghai 200032, China; ²State Key Laboratory of Oncogene and Related Genes & Department of Epidemiology, Shanghai Cancer Institute, Renji Hospital, Shanghai Jiaotong University School of Medicine, Shanghai 200032, China; ³Cancer Surveillance Section, International Agency for Research on Cancer, 150 Cours Albert Thomas, F-69372 Lyon Cedex 08, France

Correspondence to: Yong-Bing Xiang. Shanghai Cancer Institute, Renji Hospital, Shanghai Jiaotong University School of Medicine, Shanghai 200032, China. Email: ybxiang@shsci.org.

Background: The temporal relationship between hepatitis B virus (HBV) mutations and hepatocellular carcinoma (HCC) remains unclear.

Methods: We conducted a meta-analysis including cohort and nested case-control studies to prospectively examine the HCC risk associated with common variants of HBV in the PreS, Enhancer II, basal core promoter (BCP) and precore regions. Pertinent studies were identified by searching PubMed, Web of Science and the Chinese Biological Medicine databases through to November 2014. Study-specific risk estimates were combined using fixed or random effects models depending on whether significant heterogeneity was detected.

Results: Twenty prospective studies were identified, which included 8 cohort and 12 nested case-control studies. There was an increased risk of HCC associated with any PreS mutations with a pooled relative risk (RR) of 3.82 [95% confidence interval (CI): 2.59-5.61]. The pooled-RR for PreS deletion was 3.98 (95% CI: 2.28-6.95), which was higher than that of PreS2 start codon mutation (pooled-RR=2.63, 95% CI: 1.30-5.34). C1653T in Enhancer II was significantly associated with HCC risk (pooled-RR=1.83; 95% CI: 1.21-2.76). For mutations in BCP, statistically significant pooled-RRs of HCC were obtained for T1753V (pooled-RR=2.09; 95% CI: 1.49-2.94) and A1762T/G1764A double mutations (pooled-RR=3.11; 95% CI: 2.08-4.64). No statistically significant association with HCC risk was observed for G1896A in the precore region (pooled-RR=0.77; 95% CI: 0.47-1.26).

Conclusions: This study demonstrated that PreS mutations, C1653T, T1753V, and A1762T/G1764A, were associated with an increased risk of HCC. Clinical practices concerning the HCC risk prediction and diagnosis may wish to focus on patients with these mutations.

Keywords: Hepatitis B virus (HBV); mutation; hepatocellular carcinoma (HCC); prospective study; meta-analysis

Introduction

Liver cancer is the fifth most frequently diagnosed cancer and the second leading cause of cancer death worldwide, accounting for almost 6% (782,000) of the total cancer incidence and over 9% (746,000) of the corresponding mortality in 2012. Nearly 80% of the registered cases occur in Eastern Asia and sub-Saharan Africa (1). The prognosis for liver cancer is very poor (overall ratio of mortality to incidence of 0.95) (1), with a 5-year relative survival of less than 20% worldwide (2-4). As such, any promotion of primary prevention, prediction and early diagnosis of liver
cancer is a critical public health concern and challenge. About 80% of hepatocellular carcinoma (HCC) cases are estimated to be associated with hepatitis B virus (HBV) infection, making it one of the most important risk factors in liver carcinogenesis (5). HBV is a partially double-stranded DNA virus that contains four overlapping open reading frames (ORFs): C, the core (nucleocapsid) protein; S, the surface protein; P, viral DNA polymerase (which contains reverse transcriptase activity); and X, the function of which is not totally clear (6). HBV DNA replicates via reverse transcription of an RNA intermediate, but as the reverse transcriptase lacks proofreading function, errors in HBV DNA replication occur at a much higher rate than for other DNA viruses (6). Numerous studies suggest that certain mutations in the HBV genome may be associated with HCC development (7-9). Those mutations, for example, include G1613A (a G-to-A substitution at nucleotide 1613) and C1653T (a C-to-T substitution at nucleotide 1653) in the Enhancer II region, T1753V [a T-to-V (C or A or G) substitution at nucleotide 1753] and the double mutation A1762T/G1764A (an A-to-T substitution at nucleotide 1762 and a G-to-A substitution at nucleotide 1764) at the basal core promoter (BCP) region, G1896A (a G-to-A substitution at nucleotide 1896) and G1899A (a G-to-A substitution at nucleotide 1899) in the precore region (7-29).

However, previous studies vary in the quality of their design and conduct. Most of the designs are hospital-based case-control or cross-sectional studies. They are only able to present a statistical relationship between exposures and the risk of HCC. Longitudinal observations suggest that HBV mutations and HCC may interact or have a reverse pathway, and thus it is possible that certain mutations may occur concurrently or following HCC occurrence (11,22,23). Such a causal bias may produce a spurious association between a specific mutation and HCC risk. As such, prospective studies with a clearly-defined temporal relationship are more convincing than retrospective studies in elucidating potentially causal links between mutation and hepatocarcinogenesis. In addition, considering the prediction value of HBV mutations on HCC occurrence, prospective evaluations could provide more reliable risk estimates, which may guide future clinical practices on HCC screening and early diagnosis. To examine these points further, we therefore conducted a meta-analysis of prospective studies (cohort and nested case-control studies, etc.) to quantitatively evaluate the association between major HBV mutations and HCC risk.

Materials and methods

Study selection

We followed standard criteria for conducting and reporting of meta-analyses of observational studies (30). A comprehensive, computerized literature search was performed using PubMed, Web of Science and the Chinese Biological Medicine through to November, 2014. We searched the databases used the combination of the Medical Subject Heading (MeSH) terms: “hepatitis B virus”, “mutation” and “hepatocellular carcinoma” and the corresponding free terms. The identified publications were reviewed independently for their relevance to the research topic by three authors (YY, JWS and LGZ). We also manually searched the reference lists of relevant publications to identify additional studies. A set of pre-specified inclusion criteria was applied during the review and discrepancies were resolved by consensus. To be included in the meta-analysis, studies had to: (I) use a prospective design; (II) report HBV mutations as the exposure of interest; (III) report HCC or liver cancer as the outcome of interest; and (IV) provide estimates of odds ratios, rate ratios, and hazard ratios with 95% confidence intervals (CIs), or the data necessary to calculate these measures.

Because HCC is a rare outcome in general population, the odds ratio in a nested case-control study is the unbiased estimate of rate ratio or hazard ratio in a cohort study. Thus, we used the relative risk (RR) as a measurement to evaluate association between the HBV mutations and HCC risk in this meta-analysis. If multiple estimates were provided, priority was given to the multivariable adjusted risk estimates which accounted for the major confounding factors in the original studies. Where more than one study was conducted in the same population, we selected either the most recent or most applicable estimates.

Data extraction

We used a standardized protocol and reporting form to abstract the following data from each publication: the first author’s name, the year of publication, the area in which the study was conducted, the duration of follow-up in the original cohort, the age of the study population, the size of the cohort/the number of controls, the number of cases, the detection method, HBV genotype, HBV mutation sites and the corresponding RRs and 95% CIs, and the covariates included for adjustment in multivariable models.
Quality assessment

To assess the study quality, a 9-star system on the basis of the Newcastle-Ottawa Scale was used in which a study was judged on 3 broad perspectives as follows: the selection of study groups, comparability of groups, and ascertainment of either the exposure or outcome of interest for case-control or cohort studies, respectively. The high-quality study was defined as a study with ≥7 awarded stars.

Statistical analysis

To examine the associations between HBV mutations and HCC risk, we used a fixed effect model to pool the study specific estimates unless significant heterogeneity was observed, then the random effect model proposed by DerSimonian and Laird was used, which considered both the within- and between-study variations (31). We conducted subgroup analyses stratified by study design (cohort or nested case-control design), study location (China or other countries), the disease status of controls/cohort [asymptomatic hepatitis B surface antigen carriers (ASC) or chronic hepatitis B (CHB)], and HBV genotypes (C or others).

Sensitivity analyses were further conducted in which one study was removed and the rest were analyzed to evaluate whether the results were affected statistically significantly. We also repeated the analyses in high-quality studies. Heterogeneity among studies was assessed with the Q and the I² statistic and results were defined as heterogeneous for P<0.10 or I²>50% (32). Small study effects (e.g., publication bias) were evaluated by visual inspection of funnel plots and formal testing by using Egger's tests (33). When there was an indication of publication bias, we used the trim and fill method to correct for funnel plot asymmetry arising from publication bias (34).

Statistical analyses were conducted using Stata version 13.0 (StataCorp LP, College Station, Texas, USA). Two-sided P<0.05 was considered statistically significant unless otherwise specified.

Results

Literature search and study characteristics

Our systematic literature search yielded a total of 20 articles which prospectively reported the associations between HBV mutations and HCC risk. A flow diagram for the search is presented in Figure 1. Of the 3,293 records identified from the three databases, 668 records were excluded because of duplicated titles and abstracts. After review of the titles and abstracts based on the pre-specified inclusion criteria, 2,569 articles were further excluded. After reviewing full text of the remaining 56 studies, 36 studies were excluded because: (I) studies did not use a prospective design (n=27); (II) studies reported HCC survival or prognosis as the sole outcome of interest (n=5); (III) no RRs or 95% CIs were reported, or there were no enough data to calculate them (n=1); or (IV) newer data were available (n=3). The reference lists of these 36 excluded studies are presented in Supplementary Document S1. Finally, we included 20 prospective studies in the analyses (10-29); and no further studies were identified on checking the reference lists of retrieved articles.

Descriptive details of the included studies are summarized in Table 1. There were a total of 1,543 cases in the 20 prospective studies. Eight studies used the entire cohort (12,14,15,19,20,25,26,28) while the remaining 12 used nested case-control designs (10,11,13,16-18,21-24,27,29). Most of the studies were conducted among Chinese populations [10 in Mainland (10,11,13,16,18,19,22-24,27), five in Taiwan area (17,20,21,25,29) and one in Hong Kong (15)], and the others in South Korea (n=2) (12,26), Japan (n=1) (14) and Untied States (n=1) (28). Most of the included studies (n=17) used direct sequencing to detect mutations (10-12,14,15,17-21,23-29), with PreS mutations, as the most frequently reported.

PreS mutations and HCC risk

The pooled-RRs of HCC for any PreS mutations are shown in Figure 2 and Table 2. All of these included studies are defined as high-quality studies. Compared with individuals without any PreS mutations, the pooled-RR was 3.82 (95% CI: 2.59-5.61) for those with the mutations. No statistical heterogeneity was detected (P_heterogeneity=0.340, I²=11.7%). The risk was higher in Chinese population (pooled-RRs=4.02; 95% CI: 2.62-6.17). Concerning the two main mutation types in PreS region, the pooled RR for PreS deletions was 3.98 (95% CI: 2.28-6.95; I²=9.2%,...
Records identified through:
1. PubMed (n=865)
2. Web of Science (n=2,246)
3. Chinese Biological Medicine (n=182)

Records after duplicates removed (n=2,625)

Records were excluded based on screening of titles and/or abstracts using general criteria (n=2,569)
1. Not a prospective design (n=27)
2. Reporting HCC survival/prognosis as the outcome of interest (n=5)
3. No RRs or 95%CIs and no enough data to calculate them (n=1)
4. Duplicate studies or updated data available (n=3)

Records obtained from titles/abstracts screening (n=56)

Records obtained from full-text screening (n=20)

Records obtained from checking reference lists of retrieved article (n=0)

Articles included in meta-analysis (n=20)
1. Cohort studies (n=8)
2. Nested case-control studies (n=12)

Figure 1 References searched and selection of studies in the meta-analysis. HCC, hepatocellular carcinoma; RR, relative risk; CI, confidence interval.

Table 1 Characteristics of studies included in the meta-analysis, 2003-2014

<table>
<thead>
<tr>
<th>References (first author, year)</th>
<th>Study area</th>
<th>Study design</th>
<th>No. of cases</th>
<th>No. of controls/cohorts</th>
<th>Status of controls</th>
<th>Mutation</th>
<th>Detection method</th>
<th>HBV genotype</th>
<th>Adjustment/matching variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Qu (a), 2014 China NCC</td>
<td>96</td>
<td>97</td>
<td>CHB</td>
<td>Pre-S mutation</td>
<td>Sequence</td>
<td>C</td>
<td>Age, smoking, alcohol drinking</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Qu (b), 2014 China NCC</td>
<td>152</td>
<td>131</td>
<td>CHB</td>
<td>C1653T, T1753V, A1762T/G1764A, G1896A</td>
<td>Sequence</td>
<td>C</td>
<td>Age, smoking, alcohol drinking</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sinn, 2012 South Korea Cohort</td>
<td>24</td>
<td>195</td>
<td>CHB</td>
<td>Pre-S mutation</td>
<td>Sequence</td>
<td>C</td>
<td>Age, sex, HBeAg, cirrhosis, HBV-DNA level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Muñoz, 2011 China NCC</td>
<td>345</td>
<td>625</td>
<td>ASC</td>
<td>A1762T/G1764A</td>
<td>Real-time polymerase chain reaction</td>
<td>NA</td>
<td>Age, cohort, screen status and screen test results</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kusakabe, 2011 Japan Cohort</td>
<td>13</td>
<td>479</td>
<td>ASC</td>
<td>C1653T, T1753V, A1762T/G1764A, G1896A</td>
<td>Sequence</td>
<td>B, C</td>
<td>Age, sex, body mass index, smoking, alcohol drinking, ALT, c-glutamyl transpeptidase, HBeAg, HBcAg, HBV-DNA, Genotype</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yuen, 2009 Hong Kong, China Cohort</td>
<td>40</td>
<td>820</td>
<td>CHB</td>
<td>A1762T/G1764A</td>
<td>Sequence</td>
<td>B, C</td>
<td>Age, gender, HBV genotype, core promoter and precore mutations, HBeAg/anti-HBe status, HBV DNA, ALT levels and LC</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1 (continued)
<table>
<thead>
<tr>
<th>References (first author, year)</th>
<th>Study area</th>
<th>Study design</th>
<th>No. of cases</th>
<th>No. of controls/cohorts</th>
<th>Status of controls</th>
<th>Mutation</th>
<th>Detection method</th>
<th>HBV genotype</th>
<th>Adjustment/matching variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yuan, 2009</td>
<td>China</td>
<td>NCC</td>
<td>50</td>
<td>106</td>
<td>ASC</td>
<td>A1762T/G1764A</td>
<td>Real-time polymerase chain reaction</td>
<td>NA</td>
<td>Age, years between blood draw and measurement of HBV DNA double mutation, neighborhood of residence at recruitment, cigarette smoking, heavy alcohol consumption, and serum concentration of retinol</td>
</tr>
<tr>
<td>Sung, 2009</td>
<td>Taiwan, China</td>
<td>NCC</td>
<td>116</td>
<td>154</td>
<td>ASC</td>
<td>A1762T/G1764A, G1896A</td>
<td>Sequence</td>
<td>A, B, C</td>
<td>OR and 95% CI calculated on the distribution of cases and controls according to the mutations.</td>
</tr>
<tr>
<td>Fang (a), 2008</td>
<td>China</td>
<td>NCC</td>
<td>33</td>
<td>33</td>
<td>ASC</td>
<td>Pre-S mutations</td>
<td>Sequence</td>
<td>B, C</td>
<td>Age, sex and status of core promoter sequence, HBeAg, Anti-HBe, genotype, ALT</td>
</tr>
<tr>
<td>Fang (b), 2008</td>
<td>China</td>
<td>Cohort</td>
<td>61</td>
<td>2258</td>
<td>ASC</td>
<td>C1653T, T1753V, A1762T/G1764A</td>
<td>Sequence</td>
<td>NA</td>
<td>Age, sex, HBeAg and abnormal ALT were not confounders in this study.</td>
</tr>
<tr>
<td>Yang, 2008</td>
<td>Taiwan, China</td>
<td>Cohort</td>
<td>153</td>
<td>2762</td>
<td>ASC</td>
<td>A1762T/G1764A, G1896A, Pre-S mutation (Unpublished data)</td>
<td>Sequence</td>
<td>B, C</td>
<td>Age, sex, cigarette smoking, alcohol drinking, ALT level, LC, HBV DNA level, HBV genotype</td>
</tr>
<tr>
<td>Chou, 2008</td>
<td>Taiwan, China</td>
<td>NCC</td>
<td>132</td>
<td>204</td>
<td>ASC</td>
<td>T1753V, A1762T/ G1764A, G1896A</td>
<td>Sequence</td>
<td>A, B, C</td>
<td>OR and 95% CI calculated on the distribution of cases and controls according to the mutations.</td>
</tr>
<tr>
<td>Cao, 2008</td>
<td>China</td>
<td>NCC</td>
<td>47</td>
<td>50</td>
<td>CHB</td>
<td>Pre S mutations</td>
<td>Other</td>
<td>NA</td>
<td>OR and 95% CI calculated on the distribution of cases and controls according to the mutations.</td>
</tr>
<tr>
<td>Guo, 2008</td>
<td>China</td>
<td>NCC</td>
<td>58</td>
<td>71</td>
<td>CHB</td>
<td>C1653T, T1753V, A1762T/G1764A</td>
<td>Sequence</td>
<td>B, C</td>
<td>OR and 95% CI calculated on the distribution of cases and controls according to the mutations.</td>
</tr>
<tr>
<td>Zhu, 2008</td>
<td>China</td>
<td>NCC</td>
<td>20</td>
<td>83</td>
<td>CHB</td>
<td>A1762T/G1764A</td>
<td>Sequence</td>
<td>C</td>
<td>OR and 95% CI calculated on the distribution of cases and controls according to the mutations.</td>
</tr>
<tr>
<td>Chen, 2007</td>
<td>Taiwan, China</td>
<td>Cohort</td>
<td>7</td>
<td>141</td>
<td>ASC, CHB</td>
<td>Pre-S deletions, G1896A</td>
<td>Sequence</td>
<td>B, C</td>
<td>Age, sex, ALT, total bilirubin, HBV-DNA, HBV genotypes</td>
</tr>
<tr>
<td>Jang, 2007</td>
<td>South Korea</td>
<td>Cohort</td>
<td>6</td>
<td>29</td>
<td>ASC, CHB, LC</td>
<td>A1762T/G1764A</td>
<td>Sequence</td>
<td>C</td>
<td>OR and 95% CI calculated on the distribution of cases and controls according to the mutations.</td>
</tr>
<tr>
<td>Zhang, 2007</td>
<td>China</td>
<td>NCC</td>
<td>32</td>
<td>32</td>
<td>CHB</td>
<td>A1762T/G1764A</td>
<td>Sequence</td>
<td>NA</td>
<td>Age sex, occupation, living environment</td>
</tr>
<tr>
<td>Tong, 2006</td>
<td>United States</td>
<td>Cohort</td>
<td>31</td>
<td>400</td>
<td>CHB</td>
<td>A1762T/G1764A, G1896A</td>
<td>Sequence</td>
<td>A, B, C</td>
<td>OR and 95% CI calculated on the distribution of cases and controls according to the mutations.</td>
</tr>
<tr>
<td>Kao, 2003</td>
<td>Taiwan, China</td>
<td>NCC</td>
<td>127</td>
<td>123</td>
<td>CHB</td>
<td>A1762T/G1764A</td>
<td>Sequence</td>
<td>B, C</td>
<td>Age, sex, LC, genotype</td>
</tr>
</tbody>
</table>

NCC, nested case-control study; ASC, asymptomatic hepatitis B surface antigen carriers; CHB, chronic hepatitis B; LC, liver cirrhosis; HBeAg, hepatitis B e antigen; ALT, alanine aminotransferase; NA, not available; other, restriction fragment length polymorphism and polyacrylamide gel electrophoresis.
pooled-RR=2.63, 95% CI: 1.30-5.34; I²=0.0%, P for heterogeneity=0.673). There was no evidence of publication bias for PreS mutations as tested using Egger’s test (P=0.139) and by visual inspection of funnel plot (Figure S1).

C1653T and HCC risk

As shown in Figure 3 and Table 2, we observed a statistically significant association between C1653T and HCC risk (pooled-RR=1.83; 95% CI: 1.21-2.76). No statistical heterogeneity was detected (P heterogeneity=0.950, I²=0.0%). Compared with studies with CHB patients as control group (pooled-RR CHB=1.75, 95% CI: 1.11-2.76), the strength of association was stronger when we used ASC as control subjects (pooled-RR ASC=2.20, 95% CI: 0.85-5.72). When we excluded one study conducted in Japan, the risk estimate for the Chinese population was similar to the overall estimate. When we excluded one low-quality study, the results did not alter materially. There was no evidence of publication bias as tested using Egger’s test (P=0.487) and by visual inspection of funnel plot (Figure S2).

T1753V and HCC risk

Figure 4 and Table 2 show the pooled-RR of HCC for T1753V. There was a significant increase in HCC risk for individuals with T1753V compared with those without the mutation (RR=2.09; 95% CI: 1.49-2.94; P heterogeneity=0.933, I²=0%). The result obtained from cohort studies (pooled-RR=1.76; 95% CI: 0.82-3.78; P heterogeneity=0.533, I²=0%) was lower than that from the nested case-control studies (pooled-RR=2.19; 95% CI: 1.50-3.19; P heterogeneity=0.890,
The risk for studies with ASC as control subjects (pooled-RR\textsubscript{ASC}=2.15; 95% CI: 1.34-3.45; \( P_{\text{heterogeneity}}=0.678, I^2=0\% \)) was slightly higher than that of CHB patients as controls (pooled-RR\textsubscript{CHB}=2.04; 95% CI: 1.26-3.31; \( P_{\text{heterogeneity}}=0.844, I^2=0\% \)). When we excluded the Japanese study, the strength of the association among the Chinese population slightly increased (pooled-RR=2.16, 95% CI: 1.53-3.05; \( P_{\text{heterogeneity}}=0.967, I^2=0\% \)). When we repeated the analysis in three high-quality studies, the pooled-RR did not change materially. Evidence of publication bias was detected using Egger’s test (\( P=0.037 \)) but not Begg’s test (\( P=0.221 \)) (Figure S3), and the overall estimate did not change when we used the trim and fill method to correct for funnel plot asymmetry arising from publication bias.

### A1762T/G1764A and HCC risk

The pooled-RR of HCC for A1762T/G1764A is shown in Figure 5 and Table 2. A total of 15 studies reported the association between A1762T/G1764A double mutations and HCC risk. Individuals with these double mutations were at higher risk of developing HCC (pooled-RR=3.11; 95% CI: 2.08-4.64; \( P_{\text{heterogeneity}}=0.001, I^2=78.1\% \)) than those without these mutations. The pooled-RR derived
from cohort studies (pooled-RR=3.27; 95% CI: 1.93-5.55; \( P_{\text{heterogeneity}}=0.089, \Gamma^2=47.6\% \)) was slightly higher than that from the nested case-control studies (pooled-RR=2.96; 95% CI: 1.69-5.19; \( P_{\text{heterogeneity}}<0.001, \Gamma^2=85.0\% \)). When we stratified by the disease status among controls, the pooled-RR with ASC as control subjects was 3.34 (95% CI: 2.01-5.57; \( P_{\text{heterogeneity}}<0.001, \Gamma^2=80.2\% \)), higher than that of CHB patients as controls (pooled-RR=2.96; 95% CI: 1.41-6.21; \( P_{\text{heterogeneity}}<0.001, \Gamma^2=80.8\% \)). Moreover, compared with other study areas (Japan, South Korea and US), the pooled-RR was higher for studies conducted in China (RR=3.00; 95% CI: 1.92-4.70; \( P_{\text{heterogeneity}}<0.001, \Gamma^2=82.6\% \)). When we stratified by study quality, the pooled-RR of 10 high-quality studies was 3.48 (95% CI: 1.90-6.37; \( P_{\text{heterogeneity}}<0.001, \Gamma^2=84.0\% \)), which was higher than that of 5 low-quality studies (pooled-RR=2.48; 95% CI: 1.88-3.28; \( P_{\text{heterogeneity}}=0.394, \Gamma^2=23.3\% \)). There was no evidence of publication bias based on Egger’s test (\( P=0.945 \)) and by visual inspection of funnel plot (Figure S4).

G1896A and HCC risk

As shown in Figure 6 and Table 2, no statistically significant association was observed between G1896A and HCC risk (pooled-RR=0.77; 95% CI: 0.47-1.26; \( P_{\text{heterogeneity}}<0.001, \Gamma^2=75.3\% \)). Of note, there was a statistically significant inverse association among Chinese population (pooled-RR=0.60, 95% CI: 0.38-0.95; \( P_{\text{heterogeneity}}=0.012, \Gamma^2=68.9\% \)). In contrast, an increased risk (pooled-RR=2.06, 95% CI: 1.09-3.89; \( P_{\text{heterogeneity}}<0.001, \Gamma^2=77.1\% \)) was observed for studies conducted in other countries (Japan and United States). The pooled-RR derived from cohort studies (pooled-RR=0.74; 95% CI: 0.21-2.62; \( P_{\text{heterogeneity}}<0.001, \Gamma^2=86.1\% \)) was slightly lower than that from the nested case-control studies (pooled-RR=0.80; 95% CI: 0.61-1.05; \( P_{\text{heterogeneity}}=0.411, \Gamma^2=0\% \)). When we stratified by study quality, the pooled-RR of 3 high-quality studies was 0.46 (95% CI: 0.14-1.56; \( P_{\text{heterogeneity}}=0.08, \Gamma^2=60.4\% \)), which was lower than that of 4 low-quality studies (pooled-RR=0.96; 95% CI: 0.63-1.45; \( P_{\text{heterogeneity}}=0.05, \Gamma^2=61.5\% \)). No indication of publication bias was found via Egger’s test (\( P=0.840 \)) and by visual inspection of funnel plot (Figure S5).

In sensitivity analyses, we sequentially excluded one study to recalculate the pooled-RRs of HCC for these HBV-specific mutations. The pooled-RRs were not altered materially in comparison with their corresponding overall estimates (data not shown).

**Discussion**

This meta-analysis systematically evaluated major HBV mutations in association with the HCC risk using prospective epidemiological studies. We found that PreS mutations, C1653T, T1753V, and A1762T/G1764A, were
statistically significantly associated with an increased risk of HCC. The strength of association ranked the highest for PreS mutations, followed by A1762T/G1764A double mutations. The summarized risk estimates were robust across subgroup and sensitivity analyses.

It is biologically plausible that certain HBV mutants in essential genes such as PreS, Enhancer II and BCP may contribute to hepatocarcinogenesis. Several specific mutations in the PreS gene may induce an unbalanced production of envelope proteins that accumulate in the endoplasmic reticulum of the hepatocytes, and may activate the endoplasmic reticulum stress signaling pathways with consequent induction of oxidative DNA damage and genomic instability (35). The cytotoxic effects exerted by the intracellular accumulation of surface proteins can contribute to liver damage, favoring the progression of the disease toward cirrhosis and the development of HCC (35). It has been proposed that several ubiquitous and liver-specific trans-regulating nuclear factors such as transcription factor Sp1, CCAAT/enhancer-binding protein and HNF4 could bind HBV at nucleotides 1653, 1753, and 1762/1764 (36-38). These HBV sites overlap with the X gene—a pivotal ORF encodes the X protein. Mutations at these sites such as C1653T, T1753V, and A1762T/G1764A may alter the binding ability of the transregulating nuclear factors which is responsible for liver carcinogenesis (38). Those mutations may also lead to amino acid alterations of X protein, and such alterations in the X protein have been suggested to transactivate oncogenes which could contribute to the hepatocarcinogenesis (39).

Three previous meta-analyses have summarized the published studies and provided pooled risk estimates for different HBV mutations sites in association with HCC risk (7-9). Compared to the previous meta-analysis which reported a 46% increased risk of HCC in HBV-infected patients with precore mutation G1896A (8), we did not observe such significant association in present meta-analysis. Of note, for studies conducted in China, G1896A was inversely associated with HCC in our study (pooled-RR=0.60, 95% CI: 0.38-0.95). Furthermore, the strength of the association for C1653T, T1753V, and A1762T/G1764A in core promoter regions all attenuated in our study in comparison with previous meta-analyses. These differences in strength of association may be due to the inclusion of large numbers of retrospective or cross-sectional studies in previous meta-analyses, potentially introducing various biases, such as inverse casual and selection biases. For example, since some mutations may occur concurrently or after HCC diagnosis, a spurious causal relationship may be observed using cross-sectional or retrospective case-control designs. Over four-fifths of all previous studies assessing the association between HBV mutations and HCC risk used a hospital-based case-control design, exposing the pooled estimates to these biases. Moreover, the heterogeneity of the risk estimates between our study and the previous meta-analyses may be partly due to differences in the study location, study population, HBV genotypes etc. For example, most of studies in our meta-analysis are from China with an epidemic of the genotypes B and C, whereas previous meta-analysis included more studies conducted in...
other population or countries with other HBV genotypes. In addition, differences in adjustment for potential confounders may also account for the heterogeneity of the results between our study and previous meta-analyses.

This study has several limitations. Firstly, most studies included in this meta-analysis were conducted in Eastern Asia, where HBV genotypes B and C are dominant (40). Thus, generalization of these results to other populations with a different distribution of HBV genotypes should be interpreted with caution. Secondly, because of the nature of the observational design and the limitations of the meta-analysis approach to eliminate potential biases inherent in the original studies, selection bias and residual confounding may distort the association between HBV mutations and HCC. For example, HBV mutation data were not available for all participants because HBV DNA could not be efficiently amplified by polymerase chain reaction if the viral load was below the detection limit. Thus, a selection bias may have been introduced given HBV mutation data were more likely to be obtained from the participants with a high viral load than those with lower viral loads. In addition, some studies eligible for inclusion did not adjust or consider viral load or acceptance of the antiviral treatments in their analyses; thus the results may be biased by the residual confounding effects of these factors. Thirdly, there was significant heterogeneity for the results for some mutation sites. The observed between-study heterogeneity may result from various sources, for example, across studies, experimental conditions or detection methods may be different, the size of cohort and the length of follow-up may vary, or other modifiers or confounders such as HBV genotype and HBV DNA load may be present. Given the small number of included studies, we failed to carry out meta-regression analyses to further explore the sources of heterogeneity. Finally, publication bias may influence the results. Although there is little evidence of publication bias across mutation sites in this meta-analysis, tests for publication bias have low statistical power, especially when the number of studies is limited.

Conclusions

This study provides evidence that PreS mutations, C1653T, T1753V, and A1762T/G1764A, are associated with an increased risk of HCC. Future studies should place emphasis on the potential mechanisms to further elucidate the possibly causal links underlying these observed associations. Clinical practice concerning HCC risk prediction and diagnosis may wish to focus on patients with these mutations.

Acknowledgements

Funding: This work was supported by the fund of the National Key Basic Research Program “973 project” (2015CB554000), and grants from the State Key Project Specialized for Infectious Diseases of China (2008ZX10002-015 and 2012ZX10002008-002).

Footnote

Conflicts of Interest: The authors have no conflict of interest to declare.

References


Document S1

The reference lists of the 36 excluded studies. Studies were excluded because: (I) studies (41-50) did not use a prospective design (n=27) (51-67); (II) studies reported HCC survival or prognosis as the sole outcome of interest (n=5) (68-72); (III) no relative risks (RRs) or 95% confidence intervals (CIs) were reported, or there were not enough data to calculate them (n=1) (73); or (IV) newer data were available (n=3) (74-76).

References

### Table S1 Methodological quality of cohort studies included in the meta-analysis*

<table>
<thead>
<tr>
<th>Reference (first author, year)</th>
<th>Representativeness of exposed cohort</th>
<th>Selection of unexposed cohort</th>
<th>Ascertainment of exposure</th>
<th>Outcome of interest not present at start of study</th>
<th>Control for important factor or additional factor†</th>
<th>Outcome assessment</th>
<th>Follow-up long enough for outcomes to occur‡</th>
<th>Adequacy of follow-up of cohorts</th>
<th>Total quality scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sinn, 2012</td>
<td>–</td>
<td>☆</td>
<td>☆</td>
<td>☆</td>
<td>☆</td>
<td>–</td>
<td>☆</td>
<td>–</td>
<td>7</td>
</tr>
<tr>
<td>Kusakabe, 2011</td>
<td>☆</td>
<td>☆</td>
<td>☆</td>
<td>☆</td>
<td>☆</td>
<td>–</td>
<td>☆</td>
<td>–</td>
<td>9</td>
</tr>
<tr>
<td>Yuen, 2009</td>
<td>–</td>
<td>☆</td>
<td>☆</td>
<td>☆</td>
<td>☆</td>
<td>–</td>
<td>☆</td>
<td>–</td>
<td>7</td>
</tr>
<tr>
<td>Fang (b), 2008</td>
<td>☆</td>
<td>☆</td>
<td>☆</td>
<td>☆</td>
<td>☆</td>
<td>–</td>
<td>☆</td>
<td>–</td>
<td>7</td>
</tr>
<tr>
<td>Yang, 2008</td>
<td>☆</td>
<td>☆</td>
<td>☆</td>
<td>☆</td>
<td>☆</td>
<td>–</td>
<td>☆</td>
<td>–</td>
<td>7</td>
</tr>
<tr>
<td>Chen, 2007</td>
<td>–</td>
<td>☆</td>
<td>☆</td>
<td>☆</td>
<td>☆</td>
<td>–</td>
<td>☆</td>
<td>–</td>
<td>8</td>
</tr>
<tr>
<td>Jang, 2007</td>
<td>–</td>
<td>☆</td>
<td>☆</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>☆</td>
<td>–</td>
<td>5</td>
</tr>
<tr>
<td>Tong, 2006</td>
<td>–</td>
<td>☆</td>
<td>☆</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>4</td>
</tr>
</tbody>
</table>

*, a study could be awarded a maximum of one star for each item except for the item Control for important factor or additional factor; †, a maximum of 2 stars could be awarded for this item. Studies that matched or controlled for age and sex received one star, whereas studies that controlled for more than 3 of other important confounders such as alcohol drinking, smoking, HBV DNA, Genotype, HBeAg, ALT, liver cirrhosis received an additional star; ‡, a cohort study with a follow-up time >10 year was assigned one star.

### Table S2 Methodological quality of nested case-control studies included in the meta-analysis*

<table>
<thead>
<tr>
<th>Reference (first author, year)</th>
<th>Adequate definition of cases</th>
<th>Representativeness of cases</th>
<th>Selection of control subjects</th>
<th>Definition of control subjects</th>
<th>Control for important factor or additional factor†</th>
<th>Exposure assessment</th>
<th>Same method of ascertainment for all subjects</th>
<th>Nonresponse rate‡</th>
<th>Total quality scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Qu (a), 2014</td>
<td>☆</td>
<td>–</td>
<td>☆</td>
<td>☆</td>
<td>☆</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>7</td>
</tr>
<tr>
<td>Qu (b), 2014</td>
<td>☆</td>
<td>–</td>
<td>☆</td>
<td>☆</td>
<td>☆</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>6</td>
</tr>
<tr>
<td>Munoz, 2011</td>
<td>☆</td>
<td>–</td>
<td>☆</td>
<td>☆</td>
<td>☆</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>7</td>
</tr>
<tr>
<td>Yuan, 2009</td>
<td>☆</td>
<td>–</td>
<td>☆</td>
<td>☆</td>
<td>☆</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>9</td>
</tr>
<tr>
<td>Sung, 2009</td>
<td>☆</td>
<td>–</td>
<td>☆</td>
<td>–</td>
<td>☆</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>5</td>
</tr>
<tr>
<td>Fang (a), 2008</td>
<td>☆</td>
<td>–</td>
<td>☆</td>
<td>☆</td>
<td>☆</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>8</td>
</tr>
<tr>
<td>Chou, 2008</td>
<td>☆</td>
<td>–</td>
<td>☆</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>5</td>
</tr>
<tr>
<td>Cao, 2008</td>
<td>☆</td>
<td>☆</td>
<td>☆</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>7</td>
</tr>
<tr>
<td>Guo, 2008</td>
<td>☆</td>
<td>☆</td>
<td>☆</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>7</td>
</tr>
<tr>
<td>Zhu, 2008</td>
<td>☆</td>
<td>☆</td>
<td>☆</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>7</td>
</tr>
<tr>
<td>Zhang, 2007</td>
<td>–</td>
<td>☆</td>
<td>☆</td>
<td>☆</td>
<td>☆</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>7</td>
</tr>
<tr>
<td>Kao, 2003</td>
<td>☆</td>
<td>☆</td>
<td>☆</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>8</td>
</tr>
</tbody>
</table>

*, a study could be awarded a maximum of one star for each item except for the item Control for important factor or additional factor; †, a maximum of 2 stars could be awarded for this item. Studies that matched or controlled for age and sex received one star, whereas studies that controlled for more than 3 of other important confounders such as alcohol drinking, smoking, HBV DNA, Genotype, HBeAg, ALT, liver cirrhosis received an additional star; ‡, a One star was assigned if there was no significant difference in the response rate between control subjects and cases by using the chi-square test (P<0.05).
**Figure S1** Funnel plot for the association between PreS mutations and hepatocellular carcinoma risk. SE, standard error; RR, relative risk.

**Figure S2** Funnel plot for the association between C1653T and hepatocellular carcinoma risk. SE, standard error; RR, relative risk.

**Figure S3** Funnel plot for the association between T1753V and hepatocellular carcinoma risk. SE, standard error; RR, relative risk.

**Figure S4** Funnel plot for the association between A1762T/G1764A and hepatocellular carcinoma risk. SE, standard error; RR, relative risk.

**Figure S5** Funnel plot for the association between G1896A and hepatocellular carcinoma risk. SE, standard error; RR, relative risk.