Predictive power of hepatitis B 1762T/1764A mutations in plasma for hepatocellular carcinoma risk in Qidong, China

Alvaro Muñoz1, Jian Guo Chen1,1, Patricia A. Egner, Melinda L. Marshall, Jamie L. Johnson, Michael F. Schneider, Jian Hua Lu1, Yuan Rong Zhu1, Jin-Bing Wang1, Tao Yang Chen1, Thomas W. Kensler and John D. Groopman*

Johns Hopkins University Bloomberg School of Public Health, 615 North Wolfe Street, Baltimore, MD 21205, USA and Qidong Liver Cancer Institute, Qidong, 226200 Jiangsu Province, People’s Republic of China

*To whom correspondence should be addressed. Department of Environmental Health Sciences, Johns Hopkins Bloomberg School of Public Health, 615 North Wolfe Street, Baltimore, MD 21205, USA. Tel: +1 410 955 3720; Fax: +1 410 955 0617; Email: jgroopma@jhsph.edu

Hepatocellular carcinoma (HCC) is a leading cause of cancer mortality with nearly 700 000 deaths occurring annually. Hepatitis B virus (HBV) is a major contributor to HCC and acquired mutations in the HBV genome may accelerate its pathogenesis. In this study, a matched case–control investigation of 345 men who died of HCC and 625 controls were nested within a cohort of male hepatitis B surface antigen (HBsAg) carriers from Qidong, China. Matched preserving odds ratios (ORs) were used as a measure of association and 95% confidence intervals (CIs) as a measure of precision. Real-time polymerase chain reaction allowed for a quantitative comparison of the levels of the HBV 1762T/1764A mutation in cases and controls. A total of 278 (81%) of the cases were positive for the HBV 1762T/1764A mutation compared with 250 (40%) of the controls. The matched preserving OR of 6.72 (95% CI: 4.66 to 9.68) strongly indicated that cases were significantly more probably to have the mutation. Plasma levels of DNA harboring the HBV mutation were on average 15-fold higher in cases compared with controls (P < 0.001). Most strikingly, the level of the mutation in the 20 controls who later developed and died of HCC were on average 274-fold higher than controls who did not develop HCC. Thus, within this cohort of HBsAg carriers at high risk of developing HCC, individuals positive for the HBV 1762T/1764A mutation at enrollment were substantially more probably to subsequently develop HCC with a higher concentration of the mutation in plasma enhancing predisposition for cancer development.

Introduction

Hepatocellular carcinoma (HCC) is a major cause of cancer morbidity and mortality in many parts of the world, including Asia and sub-Saharan Africa, where there are upwards of 1 000 000 new cases each year and >370 000 deaths annually in the People’s Republic of China alone (1,2). The major etiological factors associated with development of HCC in these regions are infection with hepatitis B virus (HBV) and/or hepatitis C virus and lifetime exposure to high levels of aflatoxin B1 (AFB1) in the diet (3,4). Detailed knowledge of the etiology of HCC has spurred many mechanistic studies to understand the pathogenesis of this often fatal disease and this knowledge is critical for translation to preventive interventions in high-risk populations (3,5).

Abbreviations: CI, confidence interval; HBsAg, hepatitis B surface antigen; HBV, hepatitis B virus; HCC, hepatocellular carcinoma; OR, odds ratio; PCR, polymerase chain reaction.

These authors contributed equally to this work.
was made by one or more of the following means: (i) surgical biopsy, (ii) elevated serum \( \alpha \)-fetoprotein (levels >100 ng/ml) with consistent clinical and radiological history, (iii) positive computerized axial tomography scan and (iv) ultrasonography with consistent clinical history. The accumulation of HCC cases in this cohort proceeded at a fairly constant rate of between 40 and 45 cases per year and the mean time from diagnosis of HCC to death was 6 months. In this present matched case–control study, 396 individuals who died of liver cancer with available samples were defined as cases (Figure 1) and the presence of the HBV double 1762T/1764A mutation using real-time PCR was ascertained from 370 (93%) of the samples. This collaboration between the Qidong Liver Cancer Institute and Johns Hopkins University has been reviewed and approved by each respective Institutional Review Board for Human Research.

**Control selection and matching criteria**

The goal was to select two controls from the 5581 member cohort for each of the 370 HBV 1762T/1764A mutation informative cases. At the first stage, up to eight controls were randomly selected from all of the potential controls (matching criteria defined below) for a given case. The first two eligible controls with samples in the plasma repository were selected to comprise the matched set for each case. For 25 of the 370 cases, either no controls were eligible or plasma samples were not available from the eligible controls yielding a total of 345 cases with either one or two matched controls on whom the presence of the HBV 1762T/1764A mutation could be determined (Figure 1). A total of two controls were available for 280 of the 345 cases; the remaining 65 cases only had one control resulting in a total of 625 \((280 + 65)\) controls. To mimic cohort studies where cases can play the role of controls for previous cases, we did not preclude cases being selected as controls for those dying at earlier dates. Likewise, as in cohort studies in which individuals are controls for all cases occurring before their exit, for our matched case–control design, an individual could be selected to be the control for more than one case. In summary, we ascribed to strict principles of incidence density sampling for our selection of controls.

Controls were individually matched to each case to account for five potentially confounding factors. First, it was essential that controls were alive on the date the case died of HCC (i.e. controls had to be part of the risk set of the case). Second, the baseline age of the selected control had to be within ± 1 year of the baseline age of the case. Third, the case and controls within a matched set were selected from the same time period of cohort enrollment: 1989 or 1992. Fourth, controls were matched to cases according to whether or not they were randomly selected to participate in a regular screening program for HBsAg testing and follow-up (24). Finally, if the case and controls were screened for HBsAg, the pattern of surface antigen results for the controls had to be equivalent to that of the case. For example, if the case was surface antigen positive at the first two time points and negative at the final three time points then potential controls for this case must also have been positive at the first two time points and negative at the final three time points. If a case was missing a result at a particular time point, the result for the control could be positive or negative (or missing).

**HBV 1762T/1764A mutation detection by real-time PCR**

DNA was analyzed for the presence of the HBV 1762T/1764A double mutation on an ABI 7300 real-time PCR system utilizing wild-type and mutant-specific fluorogenic probes (Applied Biosystems, Foster City, CA) with some modifications as described previously (13,23). Briefly, DNA was isolated using a Qiaamp DNA mini-kit column (Qiagen, Valencia, CA) from 200 \( \mu l \) serum. Real-time PCR was done in a total volume of 25 \( \mu l \) containing 12.5 \( \mu l \) 2 Taqman Universal Master Mix (Applied Biosystems), 2.25 \( \mu l \) of forward and reverse primers at 10 \( \mu M \), 5 \( \mu l \) water and 2 \( \mu l \) of serum DNA or plasmid standard. Primer pair and wild-type and mutant-specific probe sequences were: HBV nts1762/1764 real-time PCR primer-forward: 5’-CCGACCTT-GAGGGCACTTCA-3’; HBV nts1762/1764 real-time PCR primer-reverse: 5’-CCATTATGCCTACGCTCCTCA-3’; HBV nts1762/1764 real-time PCR wild-type probe: VIC-AGTTAAAGGTTCTTGTAC; HBV nts1762/1764 real-time PCR mutant probe: 6FAM-AGTTAAAGGTTCTTGTAC. Thermal cycling conditions were 50°C for 2 min, 95°C for 5 min, followed by 45 cycles of 95°C for 30 s and 60°C for 1 min. Samples were run in duplicate. Each 96-well plate analyzed included serial dilutions of the wild-type and mutant plasmid standards and no DNA added controls. The threshold cycle number was determined using Applied Biosystems 7300 SDS software (Version 1.4.0.25) and was determined to be 42.

**Statistical analyses**

The unit of analysis was a matched set containing up to two controls per case. With our \( 1:m (m = 1 \) or 2) matched case–control study, the OR serves as the measure of association between presence of the HBV 1762T/1764A mutation and dying of HCC. Specifically, let \( a = 1 \) and \( b = 0 \) if the case in the matched set had the HBV mutation; otherwise let \( a = 0 \) and \( b = 1 \). Let \( c \) be the number of controls in matched set who had the HBV mutation, \( d \) be the number of controls in matched set who did not have the HBV mutation and \( m = c + d \) be number of controls in matched set. The contribution to the OR for each matched set in which the case had the HBV mutation is calculated by dividing \( c \) by \( m \). The contribution to the OR for each matched set in which the case did not have the HBV mutation is calculated by dividing \( d \) by \( m \). For the calculation of the OR, the sum of the \( c/(1 + m) \) among the matched sets where the case had the HBV mutation yielded the numerator and the sum of the \( c/(1 + m) \) among the matched sets where the case did not have the mutation yielded the denominator.

Congruent with the matched case–control design, conditional logistic regression models were used to obtain the OR and 95% CI of dying of HCC for different levels of each explanatory variable. Although the matched design precluded the assessment of the main effects of age, cohort, screening status and screening test results on the occurrence of death due to HCC (all part of the matching criteria), the design does allow for estimation of their putative effect modification (e.g. does the age at which the case dies modify the effect of the HBV mutation on HCC death?). In particular, the significance of the age at cases’ death (≤45, >45 to ≤55 or >55 years), cohort (1989 versus 1992), screening program (yes versus no), the percentage of hepatitis B surface antigen tests that were positive among those screened (100 versus <100%) and the time from screening to death for the cases (<1.5, >1.5 to 3.0, >3.0 to 5.0, >5.0 to 9.0 or >9.0 years) on the relationship between presence of HBV mutation and death from HCC were assessed.

For all cases and controls with the HBV 1762T/1764A mutation, box–percentile plots (25) were used to display the distributions of the number of PCR cycles needed to detect the mutation. Furthermore, capitalizing on the fact that some of the controls later developed HCC, the cycles needed to detect the HBV 1762T/1764A mutation of those controls were compared with those who remained alive and free of HCC. A Wilcoxon rank sum test was used to test for differences in the distribution of the number of PCR cycles needed to detect the HBV mutation in different groups.

**Results**

**Characteristics of cases and controls**

By study design, controls were similar to cases with respect to time of cohort enrollment (70% were identified in 1989), participation in a HBV screening program (68% were screened) (24), pattern of HBsAg results (89% of those participating in the HBsAg screening program were positive at each 6 month screening period) and age at baseline (median = 41.0; interquartile range = 35.2–49.4). Cases had a median time of 5.3 years (interquartile range = 2.3–8.5) from study entry until death, but as expected by study design, controls were on average 8.5 years older than the cases at the end of follow-up.

**HBV mutation status in cases and controls**

A total of 278 (81%) of these cases were positive for the HBV 1762T/1764A mutation compared with 250 (40%) of controls. Table I depicts the distribution of HBV mutation status in cases and controls.
the presence of the HBV 1762T/1764A mutation in controls according to whether or not the case had the mutation. Specifically, the left hand side of the table shows the data for the 278 matched sets in which the case had the mutation while the right hand side shows the data for the 67 matched sets in which the case did not have the mutation. The shaded rows in Table I correspond to the matched sets with one control.

In Table I, for the 278 matched sets in which cases have the HBV 1762T/1764A mutation, larger values in the column labeled \(d/(1 + m)\) indicate a greater risk of having the mutation in cases compared with controls. For instance, there were 90 matched sets in which the case had the mutation and the two controls did not (e.g. higher risk in cases). In contrast, there were 42 instances in which the case with the mutation was matched to two controls who also had the mutation (e.g. equal risk in cases and controls). The numerator of the matched set OR, 106.33 was obtained by summing the product of the frequency of occurrence of each matched set (e.g. 5 in the example above) and \(d/(1 + m)\) (2/3 and 0/3 in the examples above) over each of the rows of Table I.

In the 67 matched sets in which cases did not have the HBV 1762T/1764A mutation (right hand side of Table I), larger values in the column labeled \(c/(1 + m)\) indicate that compared with controls, cases were less probably to have the mutation. For example, there were seven matched sets in which the case did not have the mutation and the two controls did (e.g. lower risk in cases). The denominator of the matched set OR, 15.83 was obtained by summing the product of the frequency of occurrence of each matched set (e.g. 5 in the example above) and \(c/(1 + m)\) (2/3 in the example above) over each of the rows of Table I.

The value of the matched set OR, 6.72 (= 106.33/15.83) indicates that cases had an odds of having the HBV 1762T/1764A mutation that was 6.72 times larger than that of the controls. The strength of the association was extremely high as evidenced by the lower bound of the 95% CI being 4.66.

We found a non-significant trend of a greater odds of the HBV 1762T/1764A mutation with the cases’ age at the time of death (Table II). Specifically, cases dying <45, >45 to ≤55 or at >55 years of age had a 5.45 (95% CI: 3.18–9.35), 7.43 (95% CI: 4.09–13.50), and 9.10 (95% CI: 4.09–20.24) higher odds of having the mutation than controls. The recruitment cohort, whether or not the cases and controls were screened for HBV, the percentage of positive HBsAg screening results and time to death from study entry did not modify the

<table>
<thead>
<tr>
<th>Table I. Presence of HBV 1762T/1764A mutation in matched case–control sets for 345 cases of liver cancer death</th>
</tr>
</thead>
<tbody>
<tr>
<td>278 (81%) matched sets with HBV mutation present in case</td>
</tr>
<tr>
<td>No. of cases with mutation (a)</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

*Weighted by the number of matched sets.

<table>
<thead>
<tr>
<th>Table II. Modification of the association between HBV mutation status and liver cancer death by age, time on study, cohort and screening</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interaction of mutation with:</td>
</tr>
<tr>
<td>Age at cases’ death, years</td>
</tr>
<tr>
<td>&lt;45</td>
</tr>
<tr>
<td>&gt;45 to &lt;55</td>
</tr>
<tr>
<td>&gt;55</td>
</tr>
<tr>
<td>Enrollment cohort</td>
</tr>
<tr>
<td>1989</td>
</tr>
<tr>
<td>1992</td>
</tr>
<tr>
<td>Participated in screening program</td>
</tr>
<tr>
<td>Yes</td>
</tr>
<tr>
<td>No</td>
</tr>
<tr>
<td>Percentage of HBV surface antigen screening results which were positive</td>
</tr>
<tr>
<td>&lt;100%</td>
</tr>
<tr>
<td>100%</td>
</tr>
<tr>
<td>Years to death from study entry</td>
</tr>
<tr>
<td>&lt;1.5</td>
</tr>
<tr>
<td>&gt;1.5 to &lt;3.0</td>
</tr>
<tr>
<td>&gt;3.0 to &lt;5.0</td>
</tr>
<tr>
<td>&gt;5.0 to &lt;9.0</td>
</tr>
<tr>
<td>&gt;9.0</td>
</tr>
</tbody>
</table>

*P-values for heterogeneity of ORs. The P-value is determined by comparing the –2 log likelihood statistic from each of the models with interaction terms to the –2 log likelihood statistic from the model containing only mutation status. The difference in the –2 log likelihood values follows a chi-square distribution with degrees of freedom equal to the difference in the number of parameters between the two models.

b228 of 235 matched sets where cases and controls were screened.
relationship between case/control status and the presence/absence of the HBV mutation.

**Real-time PCR analysis of HBV 1762T/1764A mutation in plasma**

The real-time PCR analysis strategy permitted, for the first time, a quantitative comparison of the matched cases and controls for levels of the HBV 1762T/1764A mutation circulating in plasma within this high-risk cohort of HBsAg carriers. Figure 2 depicts the distribution of the number of PCR cycles needed to detect the mutation separately for the 278 cases and 250 controls with informative data for levels of HBV mutations in plasma. Each cycle represents a 2-fold change in the level of mutation detected. While the lower tail of both distributions were similar up to the first quartile (27.7 in controls and 27.0 in cases), the median (34.0 versus 30.1), third quartile (37.8 versus 34.0) and 90th percentile (40.1 versus 37.5) values were much higher, hence the HBV 1762T/1764A mutation was at lower levels in controls. These differences indicate that the level of the HBV mutation was 15-fold greater at the 50th percentile in cases than in controls ($P < 0.001$) since fewer cycles were needed for detection.

Figure 3 shows the box–percentile distribution of the number of PCR cycles needed to detect the HBV mutation in controls stratified by whether or not they were later diagnosed with HCC. Twenty (6%) of the 250 controls were found to subsequently develop HCC. It is readily apparent that it typically took far fewer cycles to detect the mutations in the 20 controls who later developed HCC. For instance, the first (22.3) and second quartiles (26.1) of the controls who later developed HCC were both less than the first quartile of controls who did not develop HCC (28.3). Seventy-five percent of the controls who later developed HCC used $\leq 34.8$ cycles to detect the HBV mutation while it took $>34.2$ cycles to detect the mutation in half of those who did not develop HCC. It is also noteworthy that the median number of PCR cycles needed to detect the HBV mutation in controls who later developed HCC was significantly less than the number of cycles needed to detect the mutation in the 278 cases (26.1 versus 30.1; $P = 0.031$). Strikingly, at the 50th percentile, those matched controls

![Fig. 2. Box–percentile plots showing the distribution of the number of cycles needed to detect HBV 1762T/1764A mutation by HCC case/control status among 278 HCC cases and 250 controls with mutation.](image1)

![Fig. 3. Box–percentile plots showing the distribution of the number of cycles needed to detect HBV 1762T/1764A mutation by HCC status at the end of follow-up among 250 controls with mutation. Eighteen controls died of HCC by the end of follow-up after their matched cases died of the disease; furthermore, two controls developed liver cancer and were alive at the end of follow-up.](image2)
who became cases had a 274-fold higher level of the HBV 1762T/1764A mutation in plasma at baseline. Thus, within this high-risk cohort by virtue of being an HBsAg carrier, there was a statistically significant larger fraction of people positive for the HBV 1762T/1764A mutation at baseline that subsequently developed HCC; further, a higher level of these mutations in plasma presaged cancer development.

Discussion

HBV is a major risk factor for the development of HCC in many parts of the world and its pathogenesis is enhanced through the biological consequences of mutations acquired following integration of parts of the HBV genome into the human hepatocyte. The HBV 1762T/1764A mutation studied in this investigation was originally characterized in HBV e antigen-negative people (26). This alteration affects the expression of both the hepatitis B e antigen since the mutation lies in the basal core promoter (BCP) and the HBV X gene (27). The HBV e antigen modulates the biology of inflammation and cirrhosis; however, the mechanism is still unclear, but there are substantial data that point to alteration of the immune surveillance system and immune tolerance in the presence and absence of this protein (27–29). The effects of this mutation on the protein produced by the X gene is also important since this protein has multiple functions in the cell including impact on cell cycle, apoptosis and DNA repair (30). Experimentally, the downstream effects of the formation of the double mutation at nucleotides 1762 and 1764 are being explored in model systems and recent data have shown that this alteration affects the binding of several transcription factors (HNF1 and HNF4) that are liver specific (31). The 1762T/1764A double mutation has also been demonstrated to increase the rate of HBV genome synthesis in cellular models by at least 2-fold (11,12,29). Furthermore, the X protein impacts autophagy in the hepatocyte and this may serve as another biological basis for viral selection and growth in the liver (32). Thus, the formation of the 1762T/1764A mutation has a significant potential to accelerate the development of HCC.

In a number of epidemiologic studies, the 1762T/1764A double mutation was found to occur more frequently in tumors of individuals infected with the genotype C strains of HBV, which is the most common genotype found in East Asian patients (33–35). Current studies in China have reported that the presence of this HBV mutation in HCC is strongly associated with the HBV subgroup genotype, C2 (36). This double mutation tracks in clinical studies of people infected by genotype C with an increased inflammatory response that becomes stronger as the progression of liver damage transits through chronic hepatitis and into a cirrhosis stage (28).

The underlying mechanism for the formation and selection of the HBV 1762T/1764A double mutation in vivo is unknown at this time. Clearly inflammatory responses in cells produce large quantities of reactive oxygen species that are known to modify DNA and induce mutations (37). Inflammatory responses also lead to the activation of a group of cytidine deaminases of the APOBEC family; this activation process induces hypermutations in this part of the HBV genome-coding region (38). Furthermore, HBV genome replication has been shown to be a very error prone process leading to increased guanine to adenine transition mutations (39). Future studies will be needed to decipher how formation of these changes in the HBV sequence occurs.

The design of new intervention paradigms should be accelerated by the validation of biomarkers for predicting an individual’s risk for HCC following HBV infection, a process which in turn is predicated upon understanding of the molecular pathways through which the virus mediates its effects (10,40–42). A recent meta-analysis of forty-three case–control and cohort studies compiled and compared the data on the relation of HBV PreS, enhancer II (EnhII), (BCP) and precore mutations to the risk of HCC; this meta-analysis found that the 1762T/1764A mutation investigated here had an overall OR of 3.79 (95% CI = 2.71–5.29) (16). Several studies published since this meta-analysis provide additional substantiation of the impact of the HBV 1762T/1764A mutation on HCC risk (43,44). Furthermore, we have expanded upon our earlier observations in a cohort study in Shanghai characterizing the greater than multiplicative interaction between HBV and aflatoxin for development of HCC (45) by focusing on those people in that cohort, who developed HCC and were HBsAg positive at baseline. When these high-risk subjects were examined for the presence of the 1762T/1764A mutation at baseline, the adjusted OR for the 1762T/1764A mutation was 2.5 (95% CI = 1.1, 9.2) (46). The substantial fraction of HBsAg carriers who apparently develop these mutations early in the carcinogenic process have also been independently reported (44,47,48). Collectively, these data point to the probable biological relevance of the 1762T/1764A mutation affecting viral protein function that in turn heightens carcinogenic risk.

In this current investigation, we continue to build upon our studies of the high-risk HBsAg carrier cohort in Qidong, China. A prior case study within this cohort demonstrated the increased hazard for HCC in subjects with measurable HBV 1762T/1764A mutation in plasma at baseline (23). In addition, these data showed that this mutation was detected up to 15 years prior to HCC diagnosis, indicating that this alteration could be an early biomarker for HCC risk. Since the prior study demonstrated that the stored plasma samples could yield DNA that could be assessed for HBV mutations; this follow-up investigation was designed to rigorously explore this impact by devising a matched case–control study from within this cohort. Since a number of published studies have established that total HBV DNA copy number reflecting viral load is a statistically significant risk factor for HCC (43,49), it was our goal to determine if quantitative plasma levels of the specific 1762T/1764A mutation at baseline, using real-time PCR analysis, were also predictive of HCC development. To our knowledge, this quantitative question had heretofore not been addressed. Using our matched nested case–control study design within the high-risk HBsAg carrier cohort, we found that matched controls of HCC cases with the HBV 1762T/1764A mutation at baseline had a 6.72 times higher risk for HCC (95% CI: 4.66–9.68) than matched controls of HCC cases without the double mutation. This matched case–control OR analysis was generated only on the basis of a detectable compared with non-detectable measure for 1762T/1764A mutation, as defined in the methods section. The real-time PCR analysis strategy permitted a quantitative comparison of the matched cases and controls for levels of the HBV 1762T/1764A mutation within this high-risk cohort of HBsAg carriers for the first time. There were 278 HCC cases and 250 matched controls with informative data for HBV mutations and the results showed that the level in plasma of the HBV mutation at baseline was 15-fold greater at the 50th percentile in HCC cases than in controls (P < 0.001). Since the controls selected from this high-risk cohort were already at heightened risk for HCC, it was not surprising that a substantial number (40%) were positive for the HBV mutation and this was consistent with prior observations (50). Furthermore, nearly a third of these subjects had very high levels of mutation at baseline (Figure 2). Twenty (6%) of the 250 controls were found to subsequently develop HCC and at the 50th percentile a 275-fold greater level of the HBV 1762T/1764A mutation in plasma at baseline was found. Thus, within this high-risk cohort by virtue of being an HBsAg carrier, there was a statistically significant higher fraction of people positive for the HBV 1762T/1764A mutation at baseline that subsequently developed HCC. Since, the HBV virus is thought to initially infect a small number of hepatocytes and following replication and virion production continue to infect neighboring liver cells, it is reasonable to assume that the entire liver cell population can become infected and eventually harbor the 1762T/1764A mutation (40). In such a scenario, the level of these HBV mutations in plasma may reflect the proportion of the liver cells that are progressing along the pathway to HCC. The level of this quantitative biomarker can then be used to identify high-risk people for prioritized enrollment in and continued evaluation of screening and preventive interventions in populations at risk for HCC.
References


Received February 5, 2011; accepted March 13, 2011