Short communication

Evolutionary perspective on hepatitis B virus with an expanded sampling strategy

Jun-Dong He, Jian-Mei Gao, Tao Shen, Robert W. Murphy, Xin-Min Yan

A Institute of Clinical and Basic Medical Sciences, Center of Clinical Molecular Biology, Key Laboratory for Birth Defects and Genetic Diseases, First People’s Hospital of Yunnan Province, Affiliated Hospital of Kunming University of Science and Technology, Kunming, China
B State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China
C Centre for Biodiversity and Conservation Biology, Royal Ontario Museum, Toronto, Canada

A R T I C L E   I N F O

Article history:
Received 13 June 2013
Received in revised form 19 September 2013
Accepted 20 September 2013
Available online 5 October 2013

Keywords:
Evolution
HBeAg
Sampling strategy

A B S T R A C T

To investigate the role hepatis B e antigen (HBeAg) plays in the evolution of hepatitis B virus (HBV), we sequenced the basic core promoter (BCP) and precore (preC) regions of 348 clones total from ten HBV Chinese patients. Eleven mutations were more frequent in HBeAg-negative patients than in HBeAg-positive patients. Further, the sequencing of dozens of variants was found to be necessary to obtain mutation profiles. Phylogenetic and median-joining network analyses suggested that variants from each patient had a single common ancestor (monophyly). Higher haplotype and nucleotide diversities were identified in HBeAg-negative patients. Analysis of dN/dS suggested that viruses experiencing a stronger immune response had lower haplotype diversity. Because HBeAg seroconversion was associated with viral diversity it served as an indicator of HBV evolution. Significantly, this study indicated a larger sampling of variants from each patient was required to understand effectively the properties of HBV.

© 2013 Elsevier B.V. All rights reserved.

Hepatitis B e antigen (HBeAg), an indicator of active viral replication, plays a role in regulating the immunity of natural infections (Liauw and Chu, 2009). Seropositivity of HBeAg is a risk factor for the development of hepatocellular carcinoma (HCC) (You et al., 2004; Yang et al., 2002). The development of antibodies to HBeAg (anti-HBe) and subsequent loss of the antigen, referred to as HBeAg seroconversion, are associated with the clinical remission of liver disease, and a lower incidence of cirrhosis and HCC in the majority of patients (Chu and Liauw, 2007a, 2007b). The serum concentration of hepatitis B virus (HBV) DNA in HBeAg seroconverters is greater than in non-seroconverters. The rate of molecular evolution appears to be higher in HBeAg-negative patients than in HBeAg-positive patients (Osiowy et al., 2006). High rates of mutations in the basic core promoter (BCP, nt1742–1849) and its adjacent precore (preC) regions drive the loss of HBeAg (Chen et al., 2006; Du et al., 2007; Jammeh et al., 2008). For example, mutation G1896A introduces a stop codon in the preC region, abolishing HBeAg expression (Carman et al., 1989). However, little is known about how HBeAg seroconversion affects the evolution of HBV. To address this question, we analyze sequences the BCP and preC regions of 348 clones from ten Chinese chronic HBV patients.

Serum samples were collected from ten patients recruited from the First People’s Hospital of Yunnan Province (Kunming). We obtained the standard panel of assays for HBV. All patients were positive for hepatitis B surface antigen (HBsAg) and hepatitis B core antibody (HBcAb), and negative for hepatitis B surface antibody (HBsAb). Among them, four patients were HBeAg seropositive and six were seronegative (Table 1). The presence of HBsAg, HBcAb, HBeAg, and HBcAb was investigated by enzyme linked immunosorbent assay (ELISA) (Kehua Bio-engineering, Shanghai, China). None of the patients was co-infected with hepatitis C, hepatitis D or human immunodeficiency virus, and none had received treatment with interferon or nucleoside analogs prior to sampling.

BCP and preC regions were amplified by PCR to yield a product of approximately 600 bp by using the following primers: C1: 5’–CGAATTCTCTGGTGCTGCTG–3’; C2: 5’–GCTCTAGAATCCTACAGAAGCTGCC–3’. PCR products were purified with Gel Extraction Mini Kit (HuaShun BioTechnologies, Shanghai, China), cloned into pUC18 vector, and then transformed into JM109 cells according to the manufacturer’s protocols. We sequenced 348 positive clones using BigDye Terminator v3.1 Cycle Sequencing Kit and an ABI 3130 DNA Analyzer (Applied Biosystems, Foster City, CA). Sequences were edited using DNASTAR software (DNASTar Inc., Madison, WI) and aligned using CLUSTALX v1.8.1 (Thompson et al., 1997) and MEGA v5.05 (Tamura et al., 2011). To avoid potential problems and errors, a stringent
quality-control procedure was utilized in the course of sample handling and data generation (Sun et al., 2006). In addition, to preclude artificial recombination triggered by sample mix-up or contamination, each sample was amplified and sequenced individually (Shen et al., 2008). All sequences were deposited in NCBI database under the accession numbers KC163804–KC164151.

Our samples had 12 high frequency nucleotide substitutions within the BCP and preC regions (Table 2). The following 11 mutations were more frequent in HBeAg-negative patients than in HBeAg-positive patients: G1658A/C, A1659C/T, A1727G/T, A1755C/T/G, A1775G, G1896A, G1899A, A1762T/G1764A, T1753C/A1762T/G1764A, A1762T/G1764A/C1766T, and T1753C/A1762T/G1764A/C1766T. In contrast, mutation T1758C was more frequent in HBeAg-positive patients than in HBeAg-negative patients. Mutation G1896A, which was previously verified to be responsible for HBeAg seroconversion (Carman et al., 1989), was observed in our HBeAg-negative patients. The sequencing results showed that HBeAg-negative patients had more nucleotide mutations than HBeAg-positive patients, which corresponded to patterns in other populations (Zheng et al., 2011; Ayari et al., 2012). In sequencing dozens of viral variants from each individual, our research provided greater information on mutation profiles. For example, the frequency of mutation G1658A/C of patient 07 was 26/32, A1659C/T was 29/32, and A1755C/T/G was 2/32.

The immune response should be the most important selective force during the natural course of HBV infection. Previous studies suggested that serum HBeAg might serve an immunoregulatory function during natural infection (Milich and Liang, 2003). Because HBeAg expression is not essential for virus replication, the virus can switch off HBeAg expression and evade anti-HBe immunity by selecting escape mutants (Tong et al., 2005). The first discovered major mutation of HBV “G1896A” truncates the protein into a 28-aa peptide (Carman et al., 1989). The common core promoter mutations, A1762T/G1764A, down regulate HBeAg mRNA production, resulting in reduced protein levels, albeit by 20% only (Parekh et al., 2003). We found mutation G1896A in patients 08 and 09 and mutations A1762T/G1764A and T1753C/A1762T/G1764A were in most HBeAg-negative patients. Mutations that emerge under the pressure of host immune response fix quickly in HBV patients, thus abolishing HBeAg expression. Accordingly, different patients may acquire the same HBeAg status via similar mutations, such as G1896A and A1762T/G1764A. Other novel mutations also occur, such as G1658A/C, A1659C/T, and G1899A. It remains unclear whether these novel mutations are responsible for or are only associated with HBe-seroconversion. Future study involving transfection experiments may clarify the situation.

We evaluated the role played by HBeAg in the molecular evolution of HBV by constructing a phylogeny of all 348 sequences plus 23 sequences in GenBank (genotypes A, B, and C) (Fig. 1). Bayesian phylogenetic analyses were carried out using Markov Chain Monte Carlo (MCMC) simulations implemented in BEAST v1.7.1 (Drummond et al., 2012). Analyses were performed under relaxed uncorrelated lognormal and relaxed uncorrelated exponential molecular clocks using the model of nucleotide substitution (GTR + G + I). Ten million generations were sufficient to obtain the convergence of parameters. The effective sample size (ESS) values for the parameters of interest exceeded 200. The maximum clade credibility (MCC) tree was obtained from summarizing the 10,000 substitution trees after removing the first 10% as burn-in using Tree Annotator v1.7.1. Nearly all clones from each patient clustered together within the tree, thus indicating an origin from a single common ancestor (monophyletic). The only exception was patient 06, whose variants scattered among the branches. The overwhelming majority of clones from patient 03 clustered together and nested within genotype B. Variants in other subjects clustered with genotype C. These observations confirmed to the
genotype results performed by using the NCBI viral genotyping tool (http://www.ncbi.nlm.nih.gov/projects/genotyping/formpage.cgi) (Table 1). The phylogeny indicated that HBeAg-negative patients harbored greater viral diversity and more complex phylogenetic patterns than did HBeAg-positive patients (e.g., patient 07 vs. patients 02, 03, and 04) (Fig. 1).

A median-joining network constructed using NETWORK v4.6 (www.fluxus-engineering.com/sharenet.htm) (Bandelt et al., 1999) further discerned the relationships among our HBV variants (Fig. 2). Most clones from each patient clustered together, as found in the phylogeny. The four HBeAg-positive patients were as follows: clones from patient 02 mostly clustered into three haplotypes (40/45, 88.9%); clones from patient 03 mainly clustered into one haplotype (27/31, 87.1%); and clones from patient 04 mainly clustered into one haplotype (27/34, 79.4%). In contrast, the 40 sequences from patient 01 presented complex topological structure in the network, forming a pattern similar to HBeAg-negative patients. This analysis resolved major haplotypes in most HBeAg-positive patients. Among the six HBeAg-negative patients, only two clustered into one specific haplotype (patient 08: 41/52, 78.8%; patient 10: 30/38, 78.9%). In contrast to the others, patient 03 had two genotypes (B and C). Due to the lack of a full-length HBV genome, we could not determine whether patient 03 was co-infected with two genotypes or had a recombinant virus. Higher values of haplotype diversity, calculated using DnaSP v5.10
(Librado and Rozas, 2009) (Table 1), were associated with the virus displaying greater topological complexity in the network (patients 01, 05, 06, 07, and 09). Nucleotide diversities were also investigated by using Arlequin v3.11 (Excoffier et al., 2005). That of patient 10 was the lowest and patient 07 the highest (Table 1). Combined with the median-joining network, variants with more complex haplotype topological structure had higher nucleotide diversities, such as in patients 06 and 07 (diversity: 2.7% and 2.9%, respectively) (Table 1; Fig. 2).

Viral quasispecies usually suffer a reduction in diversity upon transmission, likely due to either host immune selection or severe genetic drift (Edwards et al., 2006; Farci et al., 2006). Mounting evidence demonstrates that the immune response of the host largely contributes to the evolution of HBV and manifestation of chronic viral disease (Jazayeri et al., 2010; Lim et al., 2007; Wang et al., 2010; Webster et al., 2004). The host immune system drives viral mutations by positively selecting those that are able to evade detection and elimination. Our analyses are consistent with this scenario; the number of variants increases after HBeAg seroconversion (Table 2). Phylogenetic and network analyses cluster together almost all samples from each subject (Figs. 1 and 2). Divergence appears to have occurred after transmission and host-defense. More rapid change occurs in HBeAg-negative patients relative to HBeAg-positive patients (Fig. 2). The host immune response and molecular evolution of HBV are closely associated, and HBeAg seropositive and seronegative carriers have considerably different immune responses to the virus.

The sensitive serological marker alanine aminotransferase (ALT) can assess liver injury, liver necroinflammatory activity, and the extent of a host’s immune response to HBV (Shi et al., 2009). Higher ALT concentrations usually indicate a more vigorous immune response against HBV, and a higher chance of HBV-DNA loss and HBeAg seroconversion (Wang et al., 2010). In our study, the mean ALT value was 130.67 in HBeAg-positive patients and 74.6 in HBeAg-negative patients (Table 1). Although higher ALT levels occur in HBeAg-positive patients than in HBeAg-negative patients, the difference between two groups was not statistically significant ($P = 0.155$, Student’s $t$ test).

The ratio of non-synonymous to synonymous substitutions ($dN/dS$) serves to estimate the extent of selective pressure acting on a protein-coding gene (van Hemert et al., 2008). To estimate the immune selection pressure on HBeAg, we calculated the rates using codeml module of PAML 4.7 (Yang, 2007) (Table 1). Intriguing, $dN/dS$ was inversely correlated with haplotype diversity ($r = -0.833$, $P = 0.005$, Spearman rank correlation test). Patients 04, 08 and 10, had $dN/dS$ values of 130.07, 338.61 and 556.67 and haplotype diversities of 50.8, 60.2 and 38.1, respectively. In patients 05, 07 and 09, values of $dN/dS$ were 0.47, 1.39e-04, and 0.29, and haplotype diversities were 91.7, 99.8 and 98.9, respectively. Thus, stronger host immune responses may result in higher selection pressures on HBV (i.e. higher $dN/dS$ values); in response low frequency HBV haplotypes disappear and haplotype diversity reduces.

HBeAg seroconversion, an immunologically mediated event, is tightly associated with the complex nature of clinical disease in HBV (Ikeda et al., 2012; Yuen and Lai, 2011). The earlier HBeAg seroconversion occurs or the shorter HBeAg-seropositive phase, the more likely a patient is to experience improved long-term outcomes, including disease remission, increased likelihood of survival, and possibility loss of HBeAg or seroclearance (Chiu and Liaw, 2007a, 2007b). Again, some mutations, such as A1762T/G1764A, can reduce HBeAg expression. Experiments with site-directed mutants revealed that core promoter mutations conferred significantly higher viral genome replication (Parekh et al., 2003). Thus, core promoter mutations can have a major impact on viral DNA replication and HBeAg expression. In our analyses, variants from HBeAg-negative patients have higher haplotype diversities and nucleotide diversities than do HBeAg-positive patients (Table 1; Fig. 2). Therefore, viruses in HBeAg seropositive and seronegative patients have different evolutionary behaviors. Because HBeAg seroconversion is associated with viral diversity, it can serve as an indicator of HBV evolution.

The results of our analyses, particularly the median-joining network, illustrate the need to sample dozens of variants from each patient. For example, 40 of 45 variants fall into three haplotypes of patient 02 while the remaining five variants form four haplotypes (Fig. 2). The sampling of fewer variants likely would have missed some or all of the four low-frequency haplotypes. Even with 45 clones, admittedly, there is no guarantee we sampled all variants. A high level of sampling is especially critical for HBeAg-negative patients because they display higher diversity. Consequently, our assessment of genetic diversity, phylogeny and network in patient 06 may be inaccurate due to a low number of samples. In patient 03, only 1 variant (1/31, 0.03%) represents genotype C. The sampling of fewer variants likely would have missed this genotype. Further, our results suggest it is necessary to consider each patient as having one population of viral variants. To effectively investigate the properties of HBV, it is necessary to sequence many viral variants. This strategy may apply to other viral infections. Thus, our work provides a sampling strategy for researchers in future studies.

**Acknowledgments**

We are grateful to all the volunteers and samplers. We thank Wei-Wei Zhou, Tai-Cheng Zhou and Zhi-Yong Yuan for technical assistance. We also thank the two anonymous referees for their valuable comments.

This work was supported by the Natural Science Foundation of Yunnan Province (Grant no. 200300172), in part by special fund of Science and Technology Development of Yunnan Province in association with Kunming Medical College (2012FB087), in part by Foundation of Yunnan Provincial Bureau of Health (2010NS004), in part by National Natural Science Foundation of China (No. 811160352), in part by a Visiting Professorship for Senior International Scientists from the Chinese Academy of Sciences, and in part by Discovery Grant 3148 from the Natural Sciences and Engineering Research Council of Canada.

**References**

in the basal core promoter, precore, and core regions of hepatitis B virus from Tunisia shows a shift in genotype prevalence. Hepatitis Monthly 12, e6191.


