Evolution of hepatitis B virus surface gene and protein among Iranian chronic carriers from different provinces

Fatemeh Ramezani1, Seyed Moayed Alavian2, Ahmadreza Sadeghi1, Abolfazl Khedive1, Leila Ghalichi1, Mehdi Norouzi1, Hadi Karimzadeh1, Reza Malekzadeh1, Ghodrat Montazeri1, Azim Nejatizadeh1, Masood Ziaee3, Farshid Abedi1, Behrooz Ataei4, Majid Yaran5, Babak Sayad6, Mohamad Hosein Somi6, Gholamreza Sarizadeh6, Ismail Sanee-Moghaddam5, Fariborz Mansour-Ghanaei7, Houshang Rafatpanah4, Hossein Keyvani8, Ebrahim Kalantari9, Mehdī Saberfrooozi10, Reza Rezaee11, Maryam Daram1, Mostafa Mahabadi1, Zahra Goodarzi1, Vahdat Poortahmasebi1, Babak Geravand12, Azam Khamseh1, Masoud Mahmoodi1, Seyed Mohammad Jazayeri1*

1Hepatitis B Molecular Laboratory-Department of Virology-School of Public Health-Tehran University of Medical Sciences, Tehran, Iran.
2Middle East Liver Diseases Center (MELD Centers), Tehran, Iran.
3Mental health research center, Iran University of Medical Sciences, Tehran, Iran.
4Digestive Disease Research Center, Tehran University of Medical Sciences, Tehran, Iran.
5Research Center for Molecular Medicine, Hormozgan University of Medical Sciences, Bandar Abbas, Iran.
6Hepatitis Research Center, Department of Internal Medicine, Faculty of Medicine, Birjand University of Medical Sciences, Birjand, Iran.
7Department of Infectious Disease, Mashhad University of Medical Sciences, Iran.
8Digestive Disease Research Center, Isfahan University of Medical Sciences, Isfahan, Iran.
9Kermanshah Liver Diseases and Hepatitis Research Center, Kermanshah, Iran.
10Liver and Gastrointestinal Disease Research Center, Tabriz University of Medical Sciences, Tabriz, Iran.
11Educational Region of Khozestan Blood Transfusion Organization, Ahvaz, Iran.
12Department of Gastroenterology, Zahedan University of Medical Sciences, Zahedan, Iran.
13Gastrointestinal & Liver Diseases Research Center, Guilan University of Medical Sciences, Rasht, Iran.
14Immunology Research Center, Mashhad University of Medical Sciences, Mashhad, Iran.
15Department of Virology, School of Medicine, Iran University of Medical Sciences, Tehran, Iran.
16Gholhakc medical laboratory Tehran, Iran.
17Gastroenterohepatology Research Center, Shiraz University of Medical Sciences, Shiraz, Iran.
18Ministry of Health and Medical Education, Deputy of Curative Affairs, Budget Administration, Tehran, Iran.

Received: February 2014, Accepted: August 2014
INTRODUCTION

Hepatitis B virus (HBV) is a hepatotropic, non-cytopathic DNA virus which has been estimated to infect one third of the global population, of whom 350 to 400 million are suffering from chronic infection. This chronicity estate is responsible for major complications of HBV infection including cirrhosis and hepatocellular carcinoma.

The HBV infection prevalence in Iran has been estimated at 2.14% to 2.6% (1-3). This prevalence in males is about 25% higher than females (2.55% vs. 2.03% rates, respectively) (1). According to a systematic review by Alavian et al, about 1.5 million people are suffering from HBV infection in Iran, of whom 15% to 40% (1) are at risk of developing cirrhosis and/or hepatocellular carcinoma (HCC) without intervention (4, 5). They also estimated that 225,000 to 600,000 Iranian chronic HBV carriers are at risk of serious health problems related to HBV infection and need immediate attention (1).

Other patients are HBsAg positive carriers that may disseminate infection to healthy people vertically or horizontally. Moreover, 51% to 56% of Iranian cirrhotic patients are HBsAg positive (6, 7). The most common routes of transmission reported in national surveys are perinatal transmission and intravenous drug abuse (8). The geographic distribution of HBV infection in Iran showed heterogeneous patterns with the highest prevalence rates (more than 3% of the population) in northeastern region of the country while central and western regions showed the lowest prevalence rates (1% to 2%) (1).

Different epidemiological studies with alternative methodologies have been taken place to find out the risk factors and the routes of transmission of hepatitis B virus in Iran. Factors predicting HBV infection included family history of hepatitis B infection, history of receiving blood transfusion, hospitalization, unsafe sex, male gender, and living in city area (3, 8-10).

A national-wide Iranian study had undertaken for HBV genotyping between 2006 and 2011 by Iranian Hepatitis Network and Hepatitis B laboratory at Tehran University of Medical sciences, 360 samples from different regions and ethnic groups. An extensive molecular analysis based on surface genes and proteins of isolates carried out as a whole which published in recent years. However, the details of differences of molecular analysis between provinces have not published yet. Therefore, the aim of this study was to summarize the data extracted from demographic, biochemical, serological/molecular epidemiology analysis among different ethnic groups of Iranian HBV chronic carriers.

ABSTRACT

**Background and Objectives:** Iranian chronic HBV carrier’s population has shown a unique pattern of genotype D distribution all around the country. The aim of this study was to explore more details of evolutionary history of carriers based on structural surface proteins from different provinces.

**Materials and Methods:** Sera obtained from 360 isolates from 12 Different regions of country were used for amplification and sequencing of surface proteins. A detailed mutational analysis was undertaken.

**Results:** The total ratio for Missense/Silent nucleotide substitutions was 0.96. Sistan and Kermanshah showed the lowest rate of evolution between provinces (P = 0.055). On the other hand, Khurasan Razavi and Khoozestan contained the highest ratio (P = 0.055). The rest of regions were laid between these two extremes. Azarbayjan and Guilan showed the highest proportion of immune epitope distribution (91.3% and 96%, respectively). Conversely, Sistan and Tehran harbored the least percentage (66.6% and 68.8%, respectively). Kermanshah province contained only 5.2%, whereas Isfahan had 54.5% of B cell epitope distribution. In terms of T helper epitopes, all provinces showed a somehow homogeneity: 22.58% (Fars) to 46.6% (Khuzestan). On the other hand, distribution of substitutions within the CTL epitopes showed a wide range of variation between 6.6% (Khuzestan) and 63% (Kermanshah).

**Conclusion:** Further to low selection pressure found in Iranian population, the variations between different regions designate random genetic drift within the surface proteins. These finding would have some applications in terms of specific antiviral regimen, design of more efficient vaccine and public health issues.

**Keyword:** Hepatitis B, Surface Proteins, HBsAg
MATERIALS AND METHODS

360 HBsAg-positive chronic carriers who were referred to the Iranian Hepatitis Network (2004-2010) were enrolled in a cross-sectional study. Different regions based on population and geographical zones were chosen to cover the whole ethnic parts of country (Table 1). The capital city, Tehran, with a population around 13 million was chosen as a versatile multi-cultural and multi-ethnic region without a definite ethnic background. All patients were known chronic carriers. They had no evident of co-infection with other hepatitis virus, human immunodeficiency virus (HIV), and they were treatment-naive. The diagnosis of chronic liver disease was made by clinical, biochemical, radiological and endoscopic criteria. 5 ml aliquots of whole blood samples were withdrawn from each participant. Serum was, aseptically, separated. All sera were referred to Hepatitis B Laboratory, School of Public Health for serological and molecular analysis. The methodology was identical for all specimens. The HB viral DNA was extracted from 200 µl of sera using Qiagen Mini Blood Kit (Qiagen, Hilden, Germany) according to manufacturer’s instruction. DNA was eluted using 100 µl of elution buffer, stored at -20°C. Polymerase chain reaction (PCR) was carried out in 100 µl of a mixture containing 5 µl of the extracted DNA using standard methods. The complete surface gene was amplified using S1, S2, S6 and S7 primers which included the region of surface gene specifying HBV genotypes/subtype (amino acid positions 122-160). Direct sequencing of surface gene was carried out (Genetic Analyzer ABI-3130 DNA Sequencer, Fostercity, CA, USA) using 2 pmol of appropriate primers: S6C and S7D for surface gene. The electropherograms were examined visually using Chromas program. Sequences of surface gene were aligned using the BioEdit Package version 7.0.9. 312 sequences were submitted to Genbank, under individual accession number (Table 1).

<table>
<thead>
<tr>
<th>Ethnic group</th>
<th>Region</th>
<th>Number of samples</th>
<th>Accession Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Azari</td>
<td>(North-west)</td>
<td>17</td>
<td>HM348619-35</td>
</tr>
<tr>
<td>Khorasan</td>
<td>(North-East)</td>
<td>37</td>
<td>GU938342-61, 63-64</td>
</tr>
<tr>
<td>Guilan</td>
<td>(North)</td>
<td>5</td>
<td>HQ008867-68</td>
</tr>
<tr>
<td>Kurdish</td>
<td>(West)</td>
<td>46</td>
<td>KC176161-72</td>
</tr>
<tr>
<td>Fars</td>
<td>(South)</td>
<td>19</td>
<td>HC176142-60</td>
</tr>
<tr>
<td>Hormozgan</td>
<td>(South)</td>
<td>17</td>
<td>GU938305-12, 14-22</td>
</tr>
<tr>
<td>Balooch</td>
<td>(South-East)</td>
<td>21</td>
<td>HM348694-714</td>
</tr>
<tr>
<td>Khoozezan</td>
<td>(South-West)</td>
<td>12</td>
<td>HM348682-93</td>
</tr>
<tr>
<td>Isfahan</td>
<td>(Center)</td>
<td>19</td>
<td>GU938323-41</td>
</tr>
<tr>
<td></td>
<td>(Center, Capital)</td>
<td>119</td>
<td>KC176076-99</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>KC176100-30</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>HM358277-99</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>HM358300-29</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>HM358335-39</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>KC176131-36</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>312</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Origin of 360HBsAg-positive sera that were used as the source for HBsAg mutational analysis (312 sequences were submitted to Genbank)
For the purpose of mutational analysis, we used a genotype D sequence isolated from New Guinea reported by Okomoto et al. (accession number AB033559). Our previous search for the best comparable sequence has led to the selection of this isolate as the best fit reference sequence for comparison alignment. The differences between this isolate and Iranian were only one and six mismatches in terms of amino acid and nucleotide, respectively. The distributions of amino acid changes were determined according to known immune epitope residues of HBsAg (11). For the purpose of evolutionary history study of isolates, previously we used Silent/Missense (dS/dN) ratio. For a more clear analysis we used Missense/Silent (dN/dS) ratio in this study.

RESULTS

Basic characteristics. 360HBsAg-positive chronic patients were enrolled in this study; all were native residents between different regions of Iran. All were chronic carriers, HBV DNA positive and treatment-naive. 247(68.61%) were male and 113 (31.38%) were female with a mean age of 36.34 ±12.42 (mean±SD). The mean ALT and viral load level were 81.95 ± 64.85 IU/L (mean±SD) and 15000 copy/mL, respectively (results not shown). 20.5%(n=74) and 68.8% (n=248) were HBeAg and anti-HBe positive, respectively. Also 9.72% (n=35) and 0.83% (n=3) were negative and positive for both markers, respectively.

Genotyping results. The results of the first large phylogenetic tree (contained all the studied-sequences) revealed that Iranian HBV isolates were of genotype D, supported by 95% bootstrap value (1,000 replicates, results not shown). Phylogenetic analysis of 360 complete surface genes could distinguish five sub-genotypes, D (100%); sub-genotypes were D1:97.7% (n = 352), D2: 0.55% (n = 2), D3: 0.27% (n = 1), D5: 0.83% (n = 3) and D8: 0.55% (n = 2). Finally, all isolates belonged to subtype ayw2 (100%) (Results not shown).

Surface protein evolution history as a whole. In total, 222 (61.6%) out of 360 patients contained at least one amino acid mutation. In all, 1104 “nucleotide mutations” occurred, of them 542 (49.1%) were missense (amino acid altering) and 562 (50.9%) were silent (no amino acid changing) (results not shown).

In addition, it was possible to identify the level of S proteins evolution between isolates by measuring the mutation frequency (Missense/Silent, dN/dS) of individual sequences. The average mutation frequency of all sequences was 0.96 according to the number of mutations per site. The total number of amino acid substitutions was 542, of which 404 (74.5%) were distributed in known immune epitopes of surface protein (Table 2). The average amino acid mutation frequency was 1.68 mutations per sample. There was a probability of 0.74% for substitution per amino acid position.

Comparison of evolution history of HBsAg between provinces. Previous results as a whole were presented elsewhere (12). The results obtained from the comparison between amino acid substitutions and demographic, biochemistry and type of substitutions among provinces showed no significant associations between the number of amino acid mutations with age, gender and the ALT levels of chronic carriers. On the other hand, strong correlations were found between the HBeAg status of the patients and the occurrence of amino acid substitutions; a significant proportion of such changes occurred in anti-HBe positive individuals. Moreover, substantial relationships were found between the occurrence of mutations within and outside of surface protein immune epitope residues, a majority were occurred in anti-HBe positive patients. Also, comparison between the distribution of amino acid mutations within the B, Th and CTL epitopes among provinces showed considerable variations.

Provinces comparison. In terms of evolutionary history of HBV among different population (Table 2), due to great divergence between results obtained for all provinces, a precise molecular analysis was carried out on individual patients in each region, instead of individual provinces. Results showed that ignoring small, non-significant differences, almost all locations showed a homogenous pattern (results not shown). The exception was non-synonymous/ synonymous (dN/dS) ratio. Sistan and Kermanshah showed the least ration of evolution between provinces (0.62 and 0.73, respectively, P = 0.055). On the other hand, Khorasan Razavi and Khoozestan contained the highest ratio (2.29 and 1.5, respectively, P = 0.055). The rest of regions were laid between these two extremes. The distribution of amino acid changes showed that of 542 mutated residues, 404 (74.5%) were occurred in immune epitopes (Table 2), a novel finding that has
not been reported for treatment/vaccine naive chronic HBV patients so far. Azarbajyan and Guilan showed the highest proportion of immune epitope distribution (91.3% and 96%, respectively). Conversely, Sistan and Tehran harbored the least percentage (66.6% and 68.8%, respectively).

112 (27.7%), 111 (27.4%) and 197 (48.7%) of mutations which clustered within immune epitopes were localized to B, T helper and CTL epitopes within the surface protein, respectively. Distribution of amino acid substitution within the B cell epitope (residue 100-160) showed that Kermanshah province contained only 5.2%, whereas Isfahan had 54.5% of such distribution. In terms of T helper epitopes, all provinces showed a somehow homogeneity; 22.58% (Fars) to 46.6% (Khuzestan). On the other hand, distribution of substitutions within the CTL epitopes showed a wide range of variation between 6.6% (Khuzestan) and 63% (Kermanshah). For more molecular details as a whole, readers refer to the previous report (12).

<table>
<thead>
<tr>
<th>Azarbajyan</th>
<th>Mazandaran</th>
<th>Fars</th>
<th>Hormozgan</th>
<th>Isfahan</th>
<th>Kermanshah</th>
<th>Southern</th>
<th>Khorasan</th>
<th>Khuzezan</th>
<th>Sistan</th>
<th>Tehran</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of samples</td>
<td>17</td>
<td>5</td>
<td>21</td>
<td>19</td>
<td>18</td>
<td>19</td>
<td>46</td>
<td>25</td>
<td>14</td>
<td>12</td>
<td>21</td>
</tr>
<tr>
<td>Missense : Silent ratio</td>
<td>0.8</td>
<td>1.52</td>
<td>0.75</td>
<td>1</td>
<td>0.73</td>
<td>0.77</td>
<td>2.29</td>
<td>1.5</td>
<td>0.62</td>
<td>1.08</td>
<td>0.96</td>
</tr>
<tr>
<td>Mean of amino acid substitutions per site (%)</td>
<td>-1.35</td>
<td>-1.5</td>
<td>-2.26</td>
<td>-1.67</td>
<td>-1.58</td>
<td>-1.04</td>
<td>-0.59</td>
<td>-4.29</td>
<td>-1.42</td>
<td>-0.71</td>
<td>-1.89</td>
</tr>
<tr>
<td>Percent of amino acid substitutions in immune epitopes</td>
<td>91.3%</td>
<td>96%</td>
<td>100%</td>
<td>72%</td>
<td>80%</td>
<td>73.3%</td>
<td>79.1%</td>
<td>83.3%</td>
<td>81.6%</td>
<td>88.2%</td>
<td>66.6%</td>
</tr>
<tr>
<td>Percent of amino acid substitutions in non-immune epitopes</td>
<td>8.6%</td>
<td>4%</td>
<td>0%</td>
<td>27.9%</td>
<td>20%</td>
<td>26.6%</td>
<td>20.8%</td>
<td>16.6%</td>
<td>18.3%</td>
<td>11.7%</td>
<td>33.3%</td>
</tr>
<tr>
<td>Percent of amino acid substitutions in B cell immune epitopes</td>
<td>14.3%</td>
<td>4.1%</td>
<td>100%</td>
<td>41.9%</td>
<td>37.5%</td>
<td>54.5%</td>
<td>5.2%</td>
<td>26.6%</td>
<td>28.5%</td>
<td>46.6%</td>
<td>50%</td>
</tr>
<tr>
<td>Percent of amino acid substitutions in T helper cell immune epitopes</td>
<td>38%</td>
<td>37.5%</td>
<td>0%</td>
<td>22.5%</td>
<td>33.3%</td>
<td>22.7%</td>
<td>31.5%</td>
<td>33.3%</td>
<td>26.5%</td>
<td>46.6%</td>
<td>40%</td>
</tr>
<tr>
<td>Percent of amino acid substitutions in CTL immune epitopes</td>
<td>47.7%</td>
<td>58.3%</td>
<td>0%</td>
<td>35.4%</td>
<td>29.1%</td>
<td>22.7%</td>
<td>63.1%</td>
<td>40%</td>
<td>44.9%</td>
<td>6.6%</td>
<td>10%</td>
</tr>
</tbody>
</table>
DISCUSSION

According to CDC classification for the worldwide prevalence of HBV, Iran has been located as intermediate area with a mean prevalence of 2.14. This article represents the molecular data based on comparison between 360 chronic active patients from 12 different parts of Iran. The proportion of amino acid mutation in present reports showed that as high as 61.6% of chronic HBV patients obtain at least one amino acid mutation within the structural surface protein. This finding is much higher than other reports obtained from genotypes B, C and D worldwide. The impact of these genotypic alterations on the pathogenesis of HBV in chronic carriers phenotypically deserves further in vitro investigations.

Our previous reports on these samples indicated that a majority of such changes occurred in anti-HBe positive cases (13-15). Transition from HBeAg positivity to anti-HBe positivity is somehow a prolonged process, during which a number of amino acid changes are proportional to the interaction between virus and host immune status. Not surprising, it seems that in the presence of host immune surveillance, the occurrence of such mutations is inevitable (14).

Comparison of evolutionary changes within surface proteins between different provinces showed that Khorasan Razavi contained the most versatile features of mutation. On the other hand South Khorasan (North-East) and Sistan-Baloochestan Provinces showed the less variable features (along with Kermanshah and Hormozgan). We already discussed that the low evolutionary history found in Sistan Baloochestan and South Khorasan (14, 16). We assume that eastern border of Iran which include these provinces contains a different history of HBV evolution compared to other parts of country. We already hypothesized that the reason for low prevalence of HBV liver disease complications in these two regions, might be related to slow rate of substitutions (regardless of being synonymous or non-synonymous) during the course of chronicity (14, 16). Obviously, we cannot exclude the same hypothesis for other regions who contained Missense/Silent (dN/dS) ratio <1 (see Table 2), however, details of molecular analysis in these parts and the nature of amino acid substitutions were distinguished Eastern part from other regions with somehow the same ratio.

One of interesting features in the present study which make it different from other published data on chronic HBV carriers is the unique pattern of amino acid distribution in CTL epitopes (47.8%) and on the other hand, a low prevalence of "a" determinant variation in the surface proteins from 61.6% carriers. Previous investigation showed a prevalence of 10.8% and 11.2% of HBsAg variation in chronic carriers, mainly located within the “a” determinant (17-19). All reported surveys from Iran showed that the only HBV genotype is D. Is there any correlation between this unique pattern of genotype and the number and the nature of distribution of amino acid changes? To answer this question, other similar investigations from other parts of worlds which contain the genotype D predominant is needed to avoid bias on recombination between genotype D and other genotypes. The best candidate would be the sequences from countries of Caucasian origin like Turkey, Pakistan and India with prevalence of genotype D being: 87%, 62% and 67%, respectively (20). We did not find similar results on comparison between available surface proteins in international databases. Therefore, we assume that Iranian genetic background (perhaps types of HLA) might be responsible for a high number of amino acid changes in surface protein despite being chronic carriers. We already hypothesized that differences in distribution of HLA antigens, or other immune genes between diverse geographical areas probably contributed to the selection of amino acid variation (21).

In keeping with circulation of HBV amongst Iranian ethnic group sit seems that after seroconversion to anti-HBe, no further selection pressure exists for substitutions in surface proteins among other genotype D-infected ethnic groups (other non-Iranian Caucasians). The same situation is true for Iranian ethnic group, as the total (dN/dS) ratio was 0.96, indicating a natural negative selection pressure. However, we believe that in the long term of progression of chronicity and as time goes by, random genetic drift is responsible for such high number (and also the distribution of amino acids). In fact, once a mutation has arisen in a population what determine the outcome (to be fixed or to be lost), is not always down to how much better or worse it is (natural selection) compared to those variations already present in the population; instead it may simply be down to chance. As random genetic drift observed most strongly in small populations, substitutions of amino acid in various regions of country with versatile features of B, Th and CTL epitopes distribution (see Table 2) designate this unique mutational profile. Does this phenomenon related to either viral (uniqueness of genotype D) or host (Iranian HLA...
types) factors need to be explored in the future, applying longitudinal, cohort studies including patients from different phases of chronicity (before and after HBeAg seroconversion).

In conclusion, the significance of unique pictures of mutational distribution despite low evolutionary rate in Iranian chronic carriers came up from the present study would have some applications in terms of specific antiviral regimen, design of more efficient vaccine and public health issues in the next future.

REFERENCES


