Hepatitis B Virus Surface Antigen (HBsAg) Mutations Are Rare but Clustered in Immune Epitopes in Chronic Carriers from Sistan-Balouchestan Province, Iran

Abolfazl Khedive MSc1, Ismail Sanei-Moghaddam MD2, Seyed Moayed Alavian MD3, Esmail Saberfar PhD4, Mehdi Norouzi PhD1, MohamadAli Judaki MSc1, Shiva Ghamari MSc1, Seyed Mohammad Jazayeri MD PhD4

Abstract

Background: Hepatitis B virus (HBV) gene and protein variations have frequently been observed in chronic patients. The aims of this study were to determine the genotypes as well as the patterns of HBsAg variations in chronically-infected patients from the south-eastern part of Iran.

Methods: Twenty- one chronic inactive HBV carriers from Sistan-Balouchestan Province (an area with a low prevalence of HBV complications such as cirrhosis and hepatocellular carcinoma [HCC]) were enrolled. The surface genes were amplified, sequenced, and subsequently aligned using international and national Iranian database.

Results: All strains belonged to genotype D, subgenotype D1, and subtype ayw2. Of all 39 mutations occurred at 31 nucleotide positions, 15 (38.5%) were missense (amino acid altering) and 24 (61.5%) were silent (no amino acid changing). At the amino acid level, 15 substitutions occurred; 10 (66.67%) were distributed in different immune epitopes, five of which (33.33%) were in B cell epitopes; four (6.67%) were distributed in T helper epitopes, and one (6.67%) occurred inside CTL epitopes.

Conclusion: A narrowly-focused immune pressure has been on the surface proteins, especially at the B cell level, led to the emergence of escape mutants in these patients that might be related to the pathogenicity of HBV chronic infection. Also, due to the negative selection imposed on HBV genome and the uniqueness of genotype D in this ethnic group, complications (cirrhosis and HCC) are lower than other published studies.

Keywords: HBV genotype, HBV genotype D, HBV immune epitopes mutations


Introduction

Hepatitis B virus (HBV) is a well-known cause of acute and chronic hepatitis, and around 400 million individuals worldwide are chronically infected with this virus. Besides, more than one million deaths from end-stage HBV liver diseases, such as decompensated liver cirrhosis and hepatocellular carcinoma (HCC) occur each year. Hence, morbidity and mortality of persistent HBV infection are major public health concerns.

Recent studies have shown that HBV surface protein (HBsAg) is more variable than what was initially thought, and amino acid exchanges are scattered over the whole molecule. These changes are classified as either “variants” (determined by host HLA amino acid arrangement over a long period) or “mutations” (arose after vaccine/drug therapy). According to the former classification, HBV genome variability can usefully be classified into at least eight families (genotypes) based on surface protein variations with a characteristic geographic distribution.2–4 Furthermore, variation within a subcomponent of the S gene within the major hydrophilic region (MHR) of HBsAg, the “a determinant”, is strongly associated with subtype variation.5 Furthermore, the emergence of HBV mutants usually occurs following vaccine and/or hepatitis B immunoglobulin (HBIG) administration, with amino acid exchanges in HBsAg6–10 and cases of infection that have been missed because of failure of current serologic assays to detect some variant forms of HBsAg11–15 Moreover, the presence of HBsAg mutants has been reported in patients with chronic HBV infection who have not received either active immunization or HBIG and it is thought that in such cases the host immune pressure alone is able to drive the selection of HBV mutants.16–18

Although Middle East countries are recognized as high-endemic areas of HBV infection (2% – 20% prevalence of HBsAg),19 data on HBV genotypic prevalence in some parts of this region is lacking. Reports by Alavian, et al.20 revealed that the prevalence of HBV in Iran ranges between 1.7% and 2.5% of the general population. However, molecular epidemiologic findings, as HBV genotype, have also remained vague in many countries which are located in the Middle East. The aims of this study were to determine the pattern of molecular variations in chronic HBV carriers and to characterize the genotypes of HBV in a south-eastern region of Iran.
Patients and Methods

Patients

Twenty-one HBsAg-positive patients who were referred to the Sistan-Balouchestan Hepatitis Center between 2006 and 2009 from seven regions based on population and geographic zones were enrolled in this cross-sectional study. The group studied consisted of chronic inactive hepatitis B carriers, all of whom were anti-HBe positive with low to moderate levels of viral load and different levels of ALT around the normal range. Moreover, all patients were negative for antibodies against hepatitis C, hepatitis D, and human immunodeficiency virus.

All patients were interviewed and examined by gastroenterologists to be evaluated considering the clinical findings and the results of the investigative work-up including liver histology, ultrasonography, and laboratory tests such as serologic, biochemical, and virologic tests. Consequently, the clinical status of each patient was determined. Informed consent was taken from all patients and the study protocol was approved by the local ethics committee. Samples were tested by enzyme-linked immune sorbent assay (ELISA) commercial kits for HBsAg detection (Diaipus, Inc. USA). Two mL of sera were taken from each patient and were stored at -80° C for further investigations.

DNA extraction and polymerase chain reaction

HBV DNA was extracted from a 200 μL of aliquot of serum using Qiagen Mini Blood Kit (Qiagen, Hilden, Germany) accord-

Figure 1. Evolutionary relationships of 21 taxa. The evolutionary history was inferred using the UPGMA method. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. Two from each known HBV genotypes were chosen for genotypic specification. Isolate AY 161161 (genotype A) was selected for outgrouping.
ing to the manufacturer’s instruction. At the final step, DNA was eluted using 100 μL of elution buffer, stored in -20 °C. Polymerase chain reaction (PCR) was carried out in 100 μL of a mixture containing 5 μL of the extracted DNA, using methods recommended by the manufacturer (HotStart Taq PCR, Qiagen, Hilden, Germany). The complete surface gene was amplified using S1, S2, S6, and S7 primers (Table 1) which included the region of surface gene specifying HBV genotypes/subtype (amino acid positions 122 – 160) as described previously. A quantity of 5 μL of the second round PCR products were analyzed by electrophoresis in 1% agarose gel, stained by ethidium bromide, and visualized under UV light.

DNA sequencing
The HBsAg subtype of the sequences was defined by substitutions in the ‘a’ determinant between codons 122 and 160 inclusive. 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.

Sequence analysis
After allocating a sequence to an HBV genotype by analysis of the S gene, the surface gene amino acid/nucleotide variations that were found were compared with a reference sequence obtained from Okamoto (1988, accession number: AB033559), HBsAg sequences from Iranian isolates obtained from GenBank and NCBI, and from our own laboratory reports. Comparing to the former, any amino acid change was defined as “variant” (host HLA-determined). With regards to the latter (Iranian database sequences), amino acid differences defined as “mutation”. Sequences have been submitted to GenBank, numbered from HM348694 to HM348714.

**Table 1.** Oligonucleotide primers used for PCR and sequencing. Base positions numbered from the EcoRI site

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence 5’3’ of Oligonucleotides</th>
<th>Base Position</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>CCT GCT GGT GGC TCC AGT TC</td>
<td>75–56</td>
<td>Sense</td>
</tr>
<tr>
<td>S2</td>
<td>CCA CAA TTC (K)TTT GAC ATA CTT TCC A (K=G/T)</td>
<td>979–1003</td>
<td>Anti-sense</td>
</tr>
<tr>
<td>S6</td>
<td>GCA CAC GGA ATT CCG AGG ACT GGG GAC CCT G</td>
<td>146–113</td>
<td>Sense</td>
</tr>
<tr>
<td>S7</td>
<td>GAC ACC AAG CTG GGT TAG GGT TTA AAT GTA TAC C</td>
<td>823–857</td>
<td>Anti-sense</td>
</tr>
</tbody>
</table>

**Table 2.** Amino acid mutations within HBsAg of patient groups. B cell, T helper, and CTL epitopes

<table>
<thead>
<tr>
<th>Sample code/Amino Acid Position</th>
<th>Wild Type</th>
<th>Th Epitope</th>
<th>CTL Epitope</th>
<th>B Epitope</th>
</tr>
</thead>
<tbody>
<tr>
<td>43 G</td>
<td>5 I</td>
<td>E</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>189 T</td>
<td>6 ---</td>
<td>---</td>
<td>---</td>
<td>R</td>
</tr>
<tr>
<td>196 W</td>
<td>8 ---</td>
<td>---</td>
<td>---</td>
<td>L</td>
</tr>
<tr>
<td>216 L</td>
<td>9 ---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>207 S</td>
<td>113 ---</td>
<td>---</td>
<td>---</td>
<td>S</td>
</tr>
<tr>
<td>105 P</td>
<td>114 ---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>112 G</td>
<td>115 ---</td>
<td>N</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>119 A</td>
<td>116 ---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>120 I</td>
<td>119 ---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>121 ---</td>
<td>120 ---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>122 ---</td>
<td>121 ---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>123 ---</td>
<td>122 ---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>124 ---</td>
<td>123 ---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>125 ---</td>
<td>124 ---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>126 ---</td>
<td>125 ---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>127 ---</td>
<td>126 ---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>128 ---</td>
<td>127 ---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>129 ---</td>
<td>128 ---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>130 ---</td>
<td>129 ---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
</tbody>
</table>

Note: Samples were arranged in accordance to the arrangement of immune epitopes. Amino acids are described by single letter code and numbered from the beginning of HBsAg. *: Stop codon.

**Phylogenetic analysis**
Phylogenetic analysis was performed and a neighbor-joining phylogenetic tree constructed using the MEGA 5 employing a Kimura distance matrix. Associations were tested by bootstrap resampling analysis using 1000 replicates. Associations with a bootstrap value of greater than 70% were deemed significant.

**Results**
Twenty- one HBsAg-positive patients infected with HBV were enrolled in this study, all of whom were native residents of Sistan-Balouchestan Province (south-east of Iran). The studied population consisted of 66.7% (n = 14) males and 33.3% (n = 7) females (results not shown). Their mean age ± SD was 24.4 ± 10.2 years. There was no significant association between the demographic features of the patients with biochemistry (ALT and AST) and viral parameters (HBeAg status and viral load) of the patients (results not shown).

**Phylogenetic analysis**
Two of each eight common HBV genotypes were chosen for comparison. A genotype A sequence (accession number:
AY161161) was chosen for out grouping. All sequences belonged to genotype D. All Iranian isolates were grouped together with a genotype D (accession number: AB033559) and a subgenotype D1 (accession number: X02496) originated from New Guinea genotype D (accession number: AB033559) and a subgenotype to genotype D. All Iranian isolates were grouped together with subtype (100%), (results are not shown).

Substitutions in comparison with reference genotype D (Okamoto, AB033559)

In general, comparing with reference sequence (Okamoto, 1988), of a total of 143 changes at the nucleotide level, 107 (74.8%) and 36 (25.2%) were silent and missense, respectively (results are not shown). At the amino acid levels, all contained A70P compared to Okamoto reference (results are not shown). We believe that this substitution was assigned as “variant” (see Material and Methods). According to the above-mentioned description, 21 out of 36 amino acid changes were variants and the other 15 changes were mutations (see below).

Analysis of genotype and subtype-dependent variation within the S gene of 21 patients (amino acid positions 122 – 160) demonstrated that the only detected subtype was D (100%) and ayw2 subtype (100%), (results are not shown).

Nucleotide and amino acid substitutions

In comparison with Iran sequences obtained from the database as well as from our unpublished data, in addition to the genotypic characterization described above, the sequences of the strains showed a few variability over the sequenced regions. In general, 39 “mutations” occurred at 31 nucleotide positions, of which, 15 (38.5%) were missense (amino acid altering) and 24 (61.5%) were silent (no amino acid changing) (results are not shown). At amino acid level, 15 substitutions occurred (Table 2). Three isolates, 120, 121, and 122 had a stop codon in position 69 and 216, respectively (results are not shown). Furthermore, it was possible to identify the level of S proteins evolution between isolates by measuring the mutation frequency of individual sequences. The average mutation frequency of all sequences was 1.86 according to the number of mutations per site (results are not shown).

Amino acid mutations within the surface protein immune epitopes

In eight (38%) patients, 10 (66.67%) out of 15 amino acid mutations occurred in different immune epitopes within surface protein, five of which (33.33%) were in B cell epitopes; four (26.27%) occurred in T helper epitopes, and one (6.67%) occurred inside CTL epitopes (Table 2). Two amino acid substitutions occurred in “a” determinant region of surface protein: P127L and Y134H, which did not allocate for a certain subtype and/or genotype. Hence, they were mutation, not variants.

Amino acid mutations within the surface protein and serologic/biochemistry status

Three stop codons were found, two in position 69 (number 120 and 121) and one at position 216 (number 122) (results are not shown). We did not find any correlation between these and other point mutations with either biochemistry or serology (HBeAg/ Anti-HBe status) of the patients (results not shown).

Discussion

It was possible to sequence the S gene of 21 HBV strains from Sistan-Balouchestan Province, south-east of Iran. Genotype D, subgenotype D1, and subtype ayw2 accounted for 100% of isolates. Published and unpublished data from our laboratory indicate that there has been an obvious uniqueness of this virus genetic pattern in Iran.20,23–25 We already hypothesized that this unique pattern of homology is related to the relatively recent distribution and circulation of HBV in Iran compared with other countries in the region.25

The overall ratio of silent to missense nucleotide mutations in all patients was 1.6. Furthermore, in eight patients with immune epitopes mutation, the ratio was 1.86. This indicated that the proportion of deduced amino acid changes of chronically-infected patients was low and a negative selection pattern had been exerted on the sequences. Conversely, occurrence of 10 amino acid mutations within the latter group indicated that they were under positive selection pressure. How to explain this paradox?

Distribution of the mutations within known surface protein immune epitopes reflects the virus-host interaction with an ongoing infection period. Being a structural protein, HBsAg is an immune target. The consequence of selection pressure posed by anti-S antibodies would be the emergence of immune escape mutations in this protein which no longer could be recognized by the host immune system. The results would be the presence of virus (and sometimes with a high viral load level) in a chronically-infected patient. The occurrence of Th and CTL epitope mutations indicates an ineffective T cell response. In this regard, it has already been shown that these responses are weak and sometimes undetectable during the chronic state of the infection.26 Additionally, one third (33.3%) of mutations occurred in B cell epitope within MHR, encompassing amino acid residues 100 – 160. This was consistent with our previous data obtained from in vitro experiments, which showed that intracellular localization of HBeAg depended on the presence of mutations in different hepatitis B core gene B cell epitope mutations. Of 26 cloned samples, HBeAg was predominantly localized in nucleus in 13 samples in remission phase (as HBeAg is a nuclear antigen) and in cytoplasm of other 13 samples with active hepatitis. All samples with cytoplasmic localization contained B cell epitope mutations. Reversion of mutant sequences with cytoplasmic expression back to the wild type by mutagenesis led to shifting back to nuclear distribution.27

We did not find any correlation between the occurrence of point mutations/stop codons and clinical status of the patients. Previous studies suggested a functional impairment of stop codon mutations and/or deletions in terms of HBV biology.26–30 Further studies using molecular cloning approaches is essential to explore the effect of such mutations on the replication efficiency status of HBV in those patients.

As regards current investigations into genotype D of HBV, HBV genotype D is able to cause more severe diseases and higher rates of drug resistance in comparison to other studied genotypes. Thakur, et al.31 reported that genotype D of HBV correlates with more severe liver disease than HBV genotype A in India, especially in young HBV-infected patients, in whom it might lead to HCC. We were not able to compare our group of patients with genotype D with other groups of HBV genotypes. However, according to the epidemiologic studies, the prevalence of cirrhosis and HCC,
major complications of chronic HBV infection are relatively low in Iran (including the studied region) and HCC is not included in the list of top ten cancers in Iranian population (regardless of role of HBV as an etiology). Moreover, the response to anti-HBV therapy (including lamivudine, adefovir, and interferon) is significantly lower in patients from Sistan-Balouchestan Province than patients from other parts of Iran, in spite of their similar HBV genotypic pattern (Alaviani, unpublished data). In the spectrum of HBV chronicity, as time goes by (and especially after HBeAg seroconversion), the accumulation of mutations in different HBV proteins occurs. Although impact of such mutations on pathogenesis of cirrhosis and HCC is not clear, the rough conclusion is that due to the HBV genome negative selection and the uniqueness of genotype D in this ethnic group, these complications are lower than other published studies. A definite conclusion needs mutational analysis of sequential samples from different stages of chronically-infected individuals, ranging from inactive carriers to HCC cases in a cohort study. In conclusion, it can be inferred that there are variations in the structural protein of HBV in chronic patients. To interpret more accurately, the allocation of such molecular variations to the clinical, serologic, and biochemical pictures needs to be explored. In this scenario, even an individual variation must be taken into account.

References