Phylogenetic Analysis of Hepatitis B Virus Genotype F Complete Genome Sequences From Chilean Patients With Chronic Infection

Mauricio Venegas, Mónica V. Alvarado-Mora, Rodrigo A. Villanueva, João R. Rebello Pinho, Flair J. Carrilho, Stephen Locarnini, Lilly Yuen, and Javier Brahm

1Gastroenterology Section, Department of Medicine, University of Chile Clinical Hospital, Santiago, Chile
2Laboratory of Tropical Gastroenterology and Hepatology, São Paulo Institute of Tropical Medicine and Department of Gastroenterology, School of Medicine, University of São Paulo, São Paulo, SP, Brazil
3Laboratory of Hepatitis Viruses, Virology Program, Institute of Biomedical Sciences, Faculty of Medicine, Santiago, Chile
4Victorian Infectious Diseases Reference Laboratory, North Melbourne, Vic., Australia

Molecular epidemiological data concerning the hepatitis B virus (HBV) in Chile are not known completely. Since the HBV genotype F is the most prevalent in the country, the goal of this study was to obtain full HBV genome sequences from patients infected chronically in order to determine their subgenotypes and the occurrence of resistance-associated mutations. Twenty-one serum samples from antiviral drug-naïve patients with chronic hepatitis B were subjected to full-length PCR amplification, and both strands of the whole genomes were fully sequenced. Phylogenetic analyses were performed along with reference sequences available from GenBank (n = 290). The sequences were aligned using Clustal X and edited in the SE-AL software. Bayesian phylogenetic analyses were conducted by Markov Chain Monte Carlo simulations (MCMC) for 10 million generations in order to obtain the substitution tree using BEAST. The sequences were also analyzed for the presence of primary drug resistance mutations using CodonCode Aligner Software. The phylogenetic analyses indicated that all sequences were found to be the HBV subgenotype F1b, clustered into four different groups, suggesting that diverse lineages of this subgenotype may be circulating within this population of Chilean patients.

INTRODUCTION

Hepatitis B virus (HBV) infection is a severe global health problem, with approximately 2 billion people infected worldwide, and with more than 350 million of them suffering from chronic hepatitis B (CHB) [Zuckerman and Zuckerman, 2000; Shepard et al., 2006]. HBV is a DNA virus of the Hepadnaviridae family, which contains a genome composed of approximately 3,200 nucleotides (nt), with four overlapping but frame-shifted open-reading frames for the P, preC/C, preS1/preS2/S, and X viral genes [Tiollais et al., 1981].

Molecular variation and sequence changes in the HBV genome over time have resulted in the emergence of at least nine genotypes. The HBV genotypes A to I are classified based on an intergroup divergence of 8% or more in their nucleotide sequence over the entire genome [Okamoto et al., 1988; Norder et al., 1994; Stuyver et al., 2000; Arauz-Ruiz et al., 2002; Yu et al., 2010]. Genotypes may influence the HBeAg

KEY WORDS: HBV subgenotype F1b; resistance mutations; Santiago, Chile

© 2011 WILEY-LISS, INC.
seroconversion rate (related to mutational patterns in the pre-core and basal core promoter (BCP) regions), and the severity of liver disease. The differences encountered in the severity and progression of HBV-associated liver disease, as well as the response to anti-viral agents, in different regions of the world are probably attributed, at least in part, to the different HBV genotypes [Mahtab et al., 2008].

The natural history of CHB can be described by several distinct phases. These phases are characterized by particular serological, biochemical, and viral marker patterns, generally accompanied by the appearance of well-defined viral genomic mutations. Such mutations include the double A1762T/G1764A BCP mutation and the G1896A pre-C stop-codon mutation, often in combination with the G1899A mutation. Furthermore, additional mutations in the BCP region that may confer increased replication efficiency for the virus have also been found [Baumert et al., 1998; Parekh et al., 2003].

The HBV genotype F (HBV/F) has been identified as the most prevalent of the HBV genotypes in Central and South America, and it is mainly found among native indigenous people from South America [Devesa et al., 2008]. Genotype F can be further divided into four subgenotypes (F1–F4), with a genetic divergence of 4.3–6.1% [McMahon, 2009]. The subgenotype F1a has been found in Alaska, El Salvador, Guatemala, Costa Rica, and Nicaragua, whereas the F1b genotype has been reported in Peru and Argentina. The HBV subgenotype F2 was found in Venezuela and Brazil, where it was initially associated with fulminant hepatitis in patients co-infected with the hepatitis Delta virus. Subgenotype F3 has been identified in Venezuela, Colombia, and Panama, and, like the HBV subgenotype F2, it is also associated with fulminant hepatitis in these regions. Finally, subgenotype F4 was reported in Argentina and Bolivia [Blitz et al., 1998; Huy et al., 2006; Devesa et al., 2008; Santos et al., 2010; Alvarado-Mora et al., 2011].

In a previous study utilizing restriction fragment length polymorphism (RFLP), the HBV/F was found to be the most prevalent in Chile (84%), whereas genotypes A, B, C, and D were found at the frequencies of 3.8%, 3.8%, 6.1%, and 2.3%, respectively [Venegas et al., 2008]. In the current report, complete genome sequences of HBV isolates from 21 Chilean patients infected chronically with HBV/F were analyzed. The results shown herein identify HBV subgenotype- and antiviral resistance-associated substitutions, but no vaccine escape mutations, within the HBV genomes circulating in Chile.

MATERIALS AND METHODS

Study Population

 Serum samples were collected between March 2005 and March 2010 from 21 patients in Chile attending the Gastroenterology Section, Clinical Hospital, University of Chile (Santiago, Chile), for routine HBV DNA detection or quantitation. This laboratory is the national reference center in Chile for the molecular diagnostics of viral hepatitis, and it processes samples from medical centers throughout the country. In the current study, however, only samples collected from Santiago Metropolitan Area were included. All of the patients were anti-HBc and HBsAg positive (MEIA AxSym, Abbott, North Chicago, IL), and HBeAg+/anti-HBe negative, as determined by a commercially available kit from mimiVIDAS (Biomerieux, Craponne, France). The viral load was determined using a COBAS® TaqMan® Hepatitis B Virus test (Roche Molecular Systems, Branchburg, NJ). Viral genotyping was carried out by polymerase chain reaction (PCR) and RFLP as previously described [Venegas et al., 2008]. Chronic infection was defined by the detection of HBsAg in two serum samples collected at least 6 months apart. Three patients were co-infected with human immunodeficiency virus (HIV) (patients HCUCH3, HCUCH15, and HCUCH21), whereas none of the patients were co-infected with hepatitis C virus. All of the patients were male and their age ranged from 10 to 77 years old (mean age = 46 years). The ethics committee of the Clinical Hospital, University of Chile, approved this study and all participating patients signed an informed consent form. The HBV clinical data and GenBank accession numbers from the patients are shown in Table I.

HBV Complete Genome Amplification

Viral DNA was extracted from 500 μl of serum using a High Pure System Viral Nucleic Acid kit (Roche Molecular Systems). Amplification of the 21 complete HBV genome was carried out as previously described (P1 and P2 primers) [Günther et al., 1995].

HBV Nucleotide Sequencing

The 3.2 kb PCR products were gel-purified using the Wizard® SV Gel and PCR Clean-Up System kit (Promega, Madison WI). Complete genomes were sequenced from both strands of the viral DNA (Macrogen, Inc., Seoul, Korea) using the primers indicated in Table II. Consensus sequences were obtained by the alignment of both sequenced strands (sense and anti-sense) using MegAlign™ software from the DNAStar package (LaserGene, Inc., Madison, WI).

Phylogenetic Analyses

In order to analyze the distribution of the different HBV/F subgenotypes in the patients, the full sequences obtained from this study were genotyped by phylogenetic reconstructions using complete HBV genome reference sequences from each genotype retrieved from Genbank (n = 290). However, since there are only a few full HBV/F genome sequences, a larger dataset comprising 111 sequences with 1,278 nucleotides of the S/POL region of all HBV/F subgenotypes was also constructed with sequences obtained from
GenBank (datasets available from the authors upon request). The two datasets of the HBV sequences were aligned using Clustal X software [Thompson et al., 1997] and edited in the SE-AL software (available at http://tree.bio.ed.ac.uk/software/seal/). In order to perform the phylogenetic analysis, the missing nucleotides were coded as “missing characters” in the nexus block. Bayesian phylogenetic analyses were carried out using Bayesian Markov Chain Monte Carlo simulations implemented in BEAST v.1.5.3 [Drummond and Rambaut, 2007]. Analysis of the HBV dataset was performed under relaxed uncorrelated lognormal and relaxed uncorrelated exponential molecular clocks using the best model of nucleotide substitution (GTR + G + I) chosen in ModelTest [Posada and Crandall, 1998], and 10 million generations were sufficient to obtain the convergence of parameters. A Maximum Clade Credibility (MCC) tree was obtained from summarizing the 10,000 substitution trees using Tree Annotator v.1.5.3 [Drummond and Rambaut, 2007].

### Detection of Antiviral Resistance Substitutions

The presence of drug resistance substitutions was determined using CodonCode Aligner Software v.3.5 (available at http://www.codoncode.com/). This program includes effective software for sequence assembly, contig editing, and mutation detection. The results were confirmed by analyzing the sequences with the SeqHepB program [Yuen et al., 2007]. A dataset with 290 complete HBV genomes was used to identify changes in the 21 patients. Firstly, the mutations associated with HBIG, anti-HBs monoclonal antibody and vaccination escape were screened using data reporting 39 relevant mutations in this region [Sitnik et al., 2004], which included sG145R. Primary antiviral drug resistance substitutions at the following positions were then screened: rtI169, rtL180, rtA181, rtT184, rtS202, rtM204, rtN236, and rtM250 [Zoulim and Locarnini, 2009]. Secondary (or compensatory) mutations were also included in the analysis, such as rtV173, as reported previously [Delaney et al., 2003; Zoulim and Locarnini, 2009]. In addition, any HBV subgenotypes were identified by the presence of specific substitutions at positions 122, 160, 127, and 140 in the S gene. Finally, BCP and pre-C mutations [Baumert et al., 1998; Parekh et al., 2003] were identified via sequence comparisons with other known sequences from different HBV genotypes.

### TABLE II. Primers Used for Sequencing the HBV Genome

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Nucleotide position</th>
<th>Sequence (5’–3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SB409</td>
<td>409–432</td>
<td>CAT CCT GCT GCT ATG CCT CAT CT TTT GCA TGT TTT GCT GAC GCA A</td>
</tr>
<tr>
<td>SB1174</td>
<td>1174–1195</td>
<td>TGC CAA GTG TTT GCT GAC GCA A</td>
</tr>
<tr>
<td>SB1821</td>
<td>1821–1841</td>
<td>TTT TTC ACC TCT GCC TAA TCA</td>
</tr>
<tr>
<td>SB2373</td>
<td>2373–2392</td>
<td>GAA GAA CTC CCT CGC CTC GC</td>
</tr>
<tr>
<td>SB3010</td>
<td>3010–3031</td>
<td>GCA AAC AAG GTA GGA GTG GGA G</td>
</tr>
<tr>
<td>ASB432</td>
<td>432–408</td>
<td>AAG ATG AGG CAT AGC AGC AGG ATG</td>
</tr>
<tr>
<td>ASB1195</td>
<td>1195–1174</td>
<td>TTT CGT CAG CAA ACA CTT GCC A</td>
</tr>
<tr>
<td>ASI1825</td>
<td>1825–1806</td>
<td>AAA AAG TTG CAT GGT GCT G</td>
</tr>
<tr>
<td>ASB2392</td>
<td>2392–2373</td>
<td>GGC AGG CCA GGG AGT TCT TC</td>
</tr>
<tr>
<td>ASB3031</td>
<td>3031–3010</td>
<td>CTC CCA CCA CTA CCT TGT TTG C</td>
</tr>
</tbody>
</table>

RESULTS

Phylogenetic Analyses

The phylogenies showed that the 21 HBV/F samples were from subgenotype F1b (Figs. 1 and 2). The sequences of the 21 samples from the current study were grouped into four groups within the F1b cluster with a higher support. Figure 1 shows the first analysis of complete genomes from HBV/F derived from Chile, but, unfortunately, due to the insufficient number of complete genome sequences of HBV/F1b in GenBank, no further inferences about their distribution were possible. Figure 2 shows a phylogenetic tree that was constructed from 1,291 base-pair sequences, comprising all of the HBV/F subgenotypes. In this analysis, the 21 sequences from the patients were compared with other sequences previously reported from Chile, and from other countries. The phylogenetic analysis showed that these sequences were not closely related to each other, suggesting that different lineages of subgenotype F1b are circulating in Chile. Additionally, because the HBV subgenotype F1b obtained in this study did not produce a single cluster, it can be argued that these sequences may have resulted from separate introductions into the community.

Detection of Antiviral Resistance Mutations

When the antiviral resistance mutations were analyzed, one naive patient (HCUCH15) was found to be infected with an HBV isolate that contained lamivudine (LMV), emtricitabine (FTC), and clevudine (INN) antiviral resistance-associated substitutions (rtV173L, rtL180M, and rtM204V). This male patient had the risk factor of having sex with men, and he was also co-infected with HIV. Both F and H genotypes have been described as having T237 and S238. Previously, P237H and N238T/D substitutions were associated

![Phylogenetic tree of HBV Genotype F in Chile](image)

*Fig. 1. The Maximum Clade Credibility (MCC) tree was estimated by a Bayesian analysis of 290 complete genome sequences of hepatitis B virus strains. The posterior probabilities of the key nodes are shown above the respective nodes. The HBV/F samples obtained from Chile (n = 21, HCUCH) were analyzed together with other strains from around the world. The clusters containing the strains of other HBV genotypes collapsed.*

*J. Med. Virol. DOI 10.1002/jmv*
Fig. 2. The Maximum Clade Credibility (MCC) tree was estimated by a Bayesian analysis for a larger dataset comprising 111 sequences with 1,278 nucleotides of the S/POL region of all of the HBV/F subgenotypes. The posterior probabilities of the key nodes are shown above the respective nodes. The HBV/F samples obtained from Chile (n = 21, HCUUC) were analyzed together with other HBV/F strains from around the world.
with resistance to Adefovir [Shaw et al., 2006]. Our data reinforce the current view that these sites are highly polymorphic and that the mutations therein are not related to drug resistance.

**HBIG/Anti-HBs Monoclonal/Vaccine-Associated Changes**

All of the samples from Chile were found to be from the subtype adw4. An analysis of the “a” determinant of HBsAg revealed that none of the isolates contained changes that would affect binding to HBIG, the anti-HBs monoclonal antibody or the vaccination-associated anti-HBs. Furthermore, examination of the HCUCH5 sequence showed that it contained a two amino acid deletion at the Pre-S1 region (codons 46 and 47). No significant changes were observed in any of the HBV DNA promoter regions. However, sequences from two patients (HCUCH16 and HCUCH21) presented BCP mutations: A1762T and G1764A. Besides this, the sequence from patient HCUCH5 presented two core promoter mutations: C1768T and T1770A. Analyses of the nucleotide sequence at position 1858 showed the presence of thymine in all patients. Finally, none of the patients presented the T1753C mutation or the Pre-C mutations.

**DISCUSSION**

This is the first study to report a detailed analysis of complete HBV genomes circulating in a population in Chile. The phylogenetic analyses presented here revealed that all of the patients’ sequences were of the HBV subgenotype F1b. In Chile, only data about HBV prevalence based on the detection of either surface antigens (HBsAg) or antibodies against the viral core protein (anti-HBc) have been published [Pereira et al., 2008, and references therein]. A recent report, based on RFLP profiles, showed that genotype F is the most prevalent genotype [Venegas et al., 2008]. Similar results were later published by others, based on partial sequencing of the HBV genome [DiLello et al., 2009]. The sequences obtained in the current study were compared to sequences previously reported from Chile, and it was possible to conclude that the HBV/F1b subgenotype distribution in this country is suggestive of a viral diversification process, since there are many viral lineages circulating within the population. Finally, since many strains are present in the country, they may have entered at different time points and/or from different origins. Unfortunately, it was not possible to estimate the time of the most recent common ancestor (TM RCA) for the subgenotype F1b in Chile. Since genotype F1b sequences are found in different and distant countries in the Americas, it is possible that this genotype was widely distributed over the continent after its introduction into different populations.

The two BCP mutations, C1768T and T1770A, both found in the HCUCH5 isolate, are known to result in enhanced viral encapsidation and replication. The effect of these mutations, leading to increased encapsidation, is mediated through enhanced core protein synthesis by the mutant virus [Baumert et al., 1998].

Variability of the HBV genome during the chronic phase of the disease determines the selection for viral-resistant strains [Zoulim and Locarnini, 2009]. Several studies have reported mutations in HBsAg that alter its antigenicity. In previous studies, it was observed that the LMV resistance mutations, rtV173L, rtL180M, and rtM204V, resulted in the reduced binding of antibodies to the neutralization domain (“a” determinant) of the HBsAg [Torresi et al., 2002; Sloan et al., 2008]. Also, the rtV173L mutation, which accompanies rtL180M and rtM204V in about 10–20% of cases during LVD use, allows improved HBV replication fitness [Delaney et al., 2003; Poordad and Chee, 2010]. Moreover, genotypic resistance to TDF has been detected in several patients with HIV-HBV co-infection, and the substitution rtA194T (plus rtL180M and rtM204V) has been associated with TDF resistance [Sheldon et al., 2005; Zoulim and Locarnini, 2009]. However, the rtA194T mutation was not found in any of the samples from Chile in the current study. Reduced sensitivity to TDF has been described in patients infected with rtA181T/V and tN236T [van-Bömmel et al., 2010], but neither codon substitutions were also found. Other rt sequence changes have been implicated in Adefovir failure, including rtP237H and rtN238T/D [Shaw et al., 2006]. In this study, it was found that the genotype F presents T237 and S238 polymorphisms. These polymorphisms are not related to antiviral resistance. However, since one treatment-naïve patient with HBV antiviral resistance mutations was identified, it is important to elucidate the occurrence of potential genotypic resistance mutations in patients before they start antiviral treatment.

In conclusion, this study describes the complete genomic analysis of HBV/F1b from Chile. This subgenotype is also the most common in Argentina and the description of the different subgenotypes found in South American countries will help to understand the spread of this viral variant throughout this continent. Since so few complete HBV/F genomes have been reported to date, this analysis also provides a useful reference point for future molecular epidemiology studies of HBV in South America.

**REFERENCES**


