Article

Molecular Characterization of Near Full-Length Genomes of Hepatitis B Virus Isolated from Predominantly HIV Infected Individuals in Botswana

Motswedi Anderson 1,2, Wonderful Tatenda Choga 1, Sikhulile Moyo 1,3, Trevor Graham Bell 4, Tshepiso Mbangwa 1,5, Bonolo Bonita Phinius 1, Lynnette Bhebhe 1, Theresa Kibirige Sebunya 2, Shahin Lockman 1,3, Richard Marlink 1,3,6, Anna Kramvis 4, Max Essex 1,3, Rosemary Mubanga Musonda 1, Jason Tory Blackard 7, and Simani Gaseitsiwe 1,3,*

1 Botswana Harvard AIDS Institute Partnership, Gaborone, Botswana; motswedi.anderson@gmail.com (M.A.);
wtchoga@gmail.com (W.T.C.); smoyo@bhp.org.bw (S.M.); mbangiwat@gmail.com (T.M.);
bphinius@gmail.com (B.B.P.); lynnettebhebhe@gmail.com (L.B.); shahin.lockman@gmail.com (S.L.);
rmarlink@globalhealth.rutgers.edu (R.M.); messex@hsph.harvard.edu (M.E.);
rmusonda@bhp.org.bw (R.M.M.)
2 Department of Biological Sciences, Faculty of Science, University of Botswana, Gaborone, Botswana;
sebunyat@gmail.com
3 Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health,
Boston, MA 02138, USA
4 Hepatitis Virus Diversity Research Unit (HVDRU), Department of Internal Medicine, School of Clinical
Medicine, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg 2050, South Africa;
TrevorGrahamBell@gmail.com (T.G.B.); Anna.Kramvis@wits.ac.za (A.K.)
5 Faculty of Allied Health Sciences, University of Botswana, Gaborone, Botswana
6 Rutgers Global Health Institute, Robert Wood Johnson Medical School, Rutgers University,
New Brunswick, NJ 08901, USA
7 College of Medicine, University of Cincinnati, Cincinnati, OH 45627, USA; jason.blackard@uc.edu
* Correspondence: sgaseitsiwe@bhp.org.bw

Received: 12 June 2018; Accepted: 3 September 2018; Published: 7 September 2018

Abstract: The World Health Organization plans to eliminate hepatitis B and C Infections by 2030. Therefore, there is a need to study and understand hepatitis B virus (HBV) epidemiology and viral evolution further, including evaluating occult (HBsAg-negative) HBV infection (OBI), given that such infections are frequently undiagnosed and rarely treated. We aimed to molecularly characterize HBV genomes from 108 individuals co-infected with human immunodeficiency virus (HIV) and chronic hepatitis B (CHB) or OBI identified from previous HIV studies conducted in Botswana from 2009 to 2012. Full-length (3.2 kb) and nearly full-length (~3 kb) genomes were amplified by nested polymerase chain reaction (PCR). Sequences from OBI participants were compared to sequences from CHB participants and GenBank references to identify OBI-unique mutations. HBV genomes from 50 (25 CHB and 25 OBI) individuals were successfully genotyped. Among OBI participants, subgenotype A1 was identified in 12 (48%), D3 in 12 (48%), and E in 1 (4%). A similar genotype distribution was observed in CHB participants. Whole HBV genome sequences from Botswana, representing OBI and CHB, were compared for the first time. There were 43 OBI-unique mutations, of which 26 were novel. Future studies using larger sample sizes and functional analysis of OBI-unique mutations are warranted.

Keywords: HBV; mutations; occult hepatitis B; chronic hepatitis B; Botswana; Africa
1. Introduction

Hepatitis B virus (HBV) remains a major global health problem, with approximately 257 million people chronically infected [1]. Viral hepatitis is now the seventh leading cause of death worldwide, with a 63% mortality increase to 1.45 million from 1990 to 2013 [2]. HBV accounts for most viral hepatitis associated deaths [2]. It is endemic in Africa and Western Pacific, with about 6% chronic infections in adults [3]. In Botswana, the prevalence of HBV is between 3.8% and 10.6% in human immunodeficiency virus (HIV)-infected individuals [4–7].

HBV is a DNA virus belonging to the family Hepadnaviridae that consists of a partially double stranded 3.2 kb genome arranged in four partially overlapping open reading frames (ORFs) [8]. The polymerase (Pol) ORF encodes for the polymerase enzyme, which is responsible for DNA priming and reverse transcription during replication [9]. The pre-surface1/pre-surface2/surface (preS1/preS2/S) ORF codes for the large, middle, and small hepatitis B surface antigen (HBsAg) proteins, respectively, which are used to diagnose HBV, contain several B and T cell epitopes, and are utilized for attaching to hepatocytes [10]. The precore/core (preC/C) ORF encodes for hepatitis e antigen (HBeAg)—an indicator for viral replication and an immune regulator [9]—and hepatitis B core antigen (HBcAg)—the capsid protein. The X ORF codes for the X trans-activator protein [11].

Hepadnaviruses are unique in that they replicate via a pregenomic RNA intermediate, which is reverse transcribed by the polymerase, an enzyme with no proof-reading capabilities [12]. This leads to nucleotide misincorporations during replication and accounts for the significant HBV heterogeneity observed within individuals [13,14]. To date, HBV has been classified into at least nine genotypes (A–I) and a putative 10th genotype J according to nucleotide divergence of >7.5% at whole genome level [15–18]. Several genotypes have been divided further into subgenotypes based on intergroup nucleotide divergence of 4–8% [18].

Routine screening for HBV is performed by detection of HBsAg. However, this approach misses occult hepatitis B infections (OBI), defined as the presence of HBV DNA in HBsAg negative participants [19]. True OBI is characterized by low HBV DNA levels (<200 IU/mL) [19]. OBI is an important clinical entity that poses diagnostic challenges and, transmissible via blood donations, mother to child transmission, close contact with infected individuals, sexual transmission, and organ transplantation [20–26]. In addition, end stage clinical complications associated with OBI include cirrhosis and hepatocellular carcinoma (HCC) [27–29]. The risk of an individual with an HBsAg-negative, HBV DNA-positive individual infection is not negligible. In a recently completed study in South Africa, it was shown that the risk of a HBsAg-negative, HBV DNA-positive individual (with or without anti-HBc) developing HCC was 5.10 (2.06–12.62) compared to a risk of 34.48 (16.26–73.13) in HBsAg-positive individuals (with or without anti-HBc), adjusted for age group, sex, HCV serostatus, country of birth, and HIV status [30].

Several mechanisms may lead to OBI. These include host factors [31,32]; coinfection with other viruses like hepatitis C virus (HCV) and HIV [33,34]; and viral diversity including point mutations, deletions, and insertions [35–37]. The majority of studies that identified occult-associated mutations focused on specific genomic regions, such as the preS1/preS2/S ORF. Mutations within this region may affect HBsAg detection and secretion [37–42]. However, mutations in the Pol, X, and preC/C ORFs may also play a role in the OBI phenotype [36,43,44].

There is a call by World Health Organization to eliminate HBV and HCV infections as a public health problem by 2030 [45]. To achieve this, robust sequence data of HBV isolated from chronic hepatitis (CHB) and OBI are necessary. However, there are very few studies that evaluated full-length OBI genomes owing to the low viremia that is characteristic of OBI [35]. Furthermore, some studies did not have controls such as CHB participants from the same population to accurately identify occult-associated mutations [46,47]. Thus, the objective of this study was to conduct robust molecular characterization of nearly full-length HBV genomes from individuals with CHB and OBI in Botswana.
2. Materials and Methods

2.1. Population

This was a cross-sectional study comprising 108 known HBV-infected individuals (both CHB and OBI) from previous studies conducted at the Botswana Harvard AIDS Institute Partnership (BHP) in Gaborone. All available HBV positive samples from various cohorts were utilized to maximize the number of genomes available for analysis. One hundred participants—72 with OBI and 28 with CHB—were included from the Botswana National Evaluation Models of HIV Care (Bomolemo) study, which enrolled HIV positive individuals initiating highly active antiretroviral therapy (HAART) between 2009 and 2012. The Bomolemo study evaluated the efficacy and tolerability of tenofovir and emtricitabine (Truvada™) as the nucleoside reverse transcriptase inhibitor (NRTI) backbone as first-line HAART for adults in Botswana. The HBV screening for the Bomolemo cohort was previously described in detail [6,48]. An additional eight CHB participants—four HIV-positive and four HIV-negative pregnant women—were included from a study comparing the effects of HIV and ARV exposure on child health and neurodevelopment (Tshipidi study) [49]. HBV screening for participants from this cohort was previously described [50].

2.2. Ethical Considerations

The study was approved by the University of Botswana Institute Review Board and the Human Research Development Committee at the Botswana Ministry of Health. The Office of Human Research Administration at the Harvard T.H. Chan School of Public Health approved the Tshipidi and Bomolemo studies. Ethics permit number: PPME 13/1811V (318).

2.3. Extraction of Plasma DNA

DNA was extracted from 1 mL of plasma using QiAamp Ultrasense virus kit according to the manufacturer’s protocol (Qiagen, Hilden, Germany). An elution volume of 30 µL was used. The extracted DNA was either used immediately for polymerase chain reaction (PCR) or stored at −70 °C until used.

2.4. Amplification and Sequencing

Full-length genomes were amplified in two fragments. A 3 kb fragment was amplified by nested PCR using PrimeSTAR® GXL DNA Polymerase kit (Takara Bio Inc., Shiga, Japan) with some modifications. The master mix was composed of 5 µL of 5X PrimeSTAR GXL Buffer containing 5 mM Mg²⁺, 0.5 µL of PrimeSTAR GXL DNA Polymerase (1.25 U/µL), 0.5 µL of 5 µm P1 primer, 0.5 µL of 5 µm P2 primer, 1.0 µL of DNA extract, and 17.5 µL of dH₂O to make up a 25 µL reaction. The primers used were P1 and P2 for first round, whereas second-round primers were P3WRS and P4WRS [51,52], Table 1. The cycling conditions were 35 cycles of denaturation at 98 °C for 10 s, annealing at 50 °C for 15 s, and extension at 68 °C for 9 min. The second-round cycling conditions were similar to those of first round except for annealing temperature set at 55 °C for 15 s and extension temperature at 68 °C for 4 min. The remaining 276 base-pair (bp) of the preC/C region was amplified using One Step superscript III (Invitrogen, Waltham, MA, USA) protocol with minor modifications. A semi-nested PCR was performed using KU1 and MA3 as first round primers and MA3 and KU2 as second-round primers [53,54], Table 1. The 3 kb and 276 bp amplicons were visualized in 1% and 2% ethidium bromide stained gels, respectively. PCR products were purified using QiAquick PCR Purification (Qiagen, Hilden, Germany), and sequencing clean-up was done using ZR DNA Sequencing Clean up Kit (Zymo, Irvine, CA, USA) according to the manufacture’s protocol. Direct sequencing was then performed on an ABI 3130xl genetic analyzer (Applied Biosystems, Foster City, CA, USA) using Big Dye sequencing chemistry. The primers used for sequencing are shown in Table 1. Sequences were submitted to National Center for Biotechnology Information (NCBI) GenBank under the accession numbers MH464807 to MH464856.
Table 1. Primers used for PCR and for sequencing.

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer Sequence</th>
<th>Coordinates from EcoRI Site</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>5′CCGGAAAGCTTGAGCTTTTCACTCTCTGAATC 3′</td>
<td>1821–1841</td>
<td>[51]</td>
</tr>
<tr>
<td>P2</td>
<td>5′CCGGAAAGCTTGAGCTTTTCACTCTCTGAATC 3′</td>
<td>1823–1806</td>
<td>[51]</td>
</tr>
<tr>
<td>P3WRS</td>
<td>5′CTACTGTCAAGGCTCCAAGC 3′</td>
<td>1823–1806</td>
<td>[52]</td>
</tr>
<tr>
<td>P4WRS</td>
<td>5′CCGGAAAGCTTGAGCTTTTCACTCTCTGAATC 3′</td>
<td>1823–1806</td>
<td>[52]</td>
</tr>
<tr>
<td>KU1</td>
<td>5′CATAGCAGACCTTCTGTGGACT 3′</td>
<td>1653–1672</td>
<td>[54]</td>
</tr>
<tr>
<td>KU2</td>
<td>5′ATACTGCAAGGACCGGCTCTG 3′</td>
<td>1679–1698</td>
<td>[54]</td>
</tr>
<tr>
<td>MA3</td>
<td>5′GAAAGAAGAGCAGAAGGAAA 3′</td>
<td>1973–1954</td>
<td>[53]</td>
</tr>
<tr>
<td>HBV Z</td>
<td>5′AGCCCTCGCCAGCAGGCTGATA 3′</td>
<td>1179–1199</td>
<td>[55]</td>
</tr>
<tr>
<td>HBV 3</td>
<td>5′CGTTGCCAAGGCTGAGGTAAGG 3′</td>
<td>2478–2455</td>
<td>[55]</td>
</tr>
<tr>
<td>HBV M</td>
<td>5′GACAGACTTGCTTAAATGGGAGCCAGGACTGATGACTGAGG 3′</td>
<td>2306–2287</td>
<td>[55]</td>
</tr>
<tr>
<td>HBV P</td>
<td>5′TCATCGGATGATCTGACTGATGACTGAGG 3′</td>
<td>1292–1311</td>
<td>[55]</td>
</tr>
<tr>
<td>HBV H</td>
<td>5′GAAAGAAGAGCAGAAGGAAA 3′</td>
<td>1973–1954</td>
<td>[53]</td>
</tr>
<tr>
<td>HBV N</td>
<td>5′ACTGAGCCAGGAGCCAGAAGGACTGAGG 3′</td>
<td>1991–1965</td>
<td>[55]</td>
</tr>
<tr>
<td>Werle AS</td>
<td>5′CGTCAGCAACAGGCTGAGG 3′</td>
<td>1175–1192</td>
<td>[56]</td>
</tr>
<tr>
<td>CoreF</td>
<td>5′GTGTGACATGCCTTCCTGCTTGCCTTGCCT 3′</td>
<td>2269–2287</td>
<td>[41]</td>
</tr>
<tr>
<td>P6</td>
<td>5′GGCAGGTCCCTAGAAGAAGAAGA 3′</td>
<td>2363–2386</td>
<td>[51]</td>
</tr>
</tbody>
</table>

The primers used for sequencing are those shown in bold.
2.5. Phylogenetic Analysis

Phylogenetic trees were constructed utilizing a Bayesian Markov chain Monte Carlo (MCMC) in the Bayesian Evolutionary Analysis by Sampling Trees (BEAST) v1.8.2 (BEAST Developers) [57] program with a chain length of 100,000,000 and sampling every 10,000 generations. The analysis utilized an uncorrelated log-normal relaxed molecular clock, the Hasegawa, Kishino, and Yano (HKY) model, and the general time reversible model with gamma distributed rates of variation among sites and a proportion of invariable sites (GTR+G+I). Tracer v1.7 (BEAST Developers) [57] was used to visualize results and confirm chain convergence. Every parameter had an effective sample size (ESS) > 500 implying sufficient sampling. Tree Annotator v1.7.3 (BEAST Developers) [57] was utilized to choose the maximum clade credibility tree after a 10% burn-in. Posterior probabilities > 90% were deemed statistically significant. Trees for subgenotype A1 and D3 sequences from this study and the respective GenBank references for the whole surface region were constructed. The S ORF was used to determine the clustering of Botswana strains relative to other African HBV sequences, because there are very few whole genome HBV sequences from Africa except for South Africa.

2.6. HBV Genotype Recombination Analysis

To evaluate potential recombination, study sequences were compared to GenBank references A-H utilizing Simplot software v3.5.1 (Ray, S.C, Baltimore, MD, USA) with a step size of 20 bp and 200 bp window size [58]. The neighbor-joining method (NJ) was used to conduct a bootscan analysis utilizing 1000 bootstrap replicates and an 80% threshold.

2.7. Analysis of Immune Selection Pressure, Signature Amino Acids, and Escape Mutations

DataMonkey was used to approximate the rates of nonsynonymous (dN) and synonymous (dS) substitutions using fixed effects likelihood (FEL) [59]. For this analysis, codons for each ORF (S, preS1, preS2, Pol, X, core) were analyzed individually according to HBV status (either CHB or OBI). Codon positions were numbered from the beginning of each open reading frame. The preC region could not be amplified from several OBI participants; hence, it was excluded from this analysis. The viral epidemiology signature pattern analysis (VESPA) was utilized to search for signature amino acids (aa) in HBV sequences isolated from OBI participant’s sequences compared to those isolated from CHB participants for each ORF based on subgenotype [60]. Escape mutations were identified using the online tool Geno2pheno available at https://hbv.geno2pheno.org/. Escape mutations previously reported in the literature were also searched for manually in BioEdit.

2.8. Mutational Analysis

Genomes of HBV were aligned with GenBank references using ClustalX v2.1 (Higgins D., Sievers F., Dineen D., Wilm A, Dublin, Ireland). The subgenotype A1 sequences were aligned with 107 full-length subgenotype A1 references, while the subgenotype D3 sequences were aligned with 85 full-length subgenotype D3 references. The references used were HBV sequences isolated from CHB participants, which were extracted from online curated GenBank alignments [61]. All subgenotypes’ A1 and D3 whole genome sequences available at that time (January 2017) were included. The sequences were then trimmed to the same length in BioEdit v7.2.5 (Hall T., Carlsbad, CA, USA). Subsequently, Babylon Translator was utilized to extract each ORF (Pre S1, Pre S2, S, X, PreC, core, Pol) and the sequences translated into aa [62]. To identify potential OBI-unique mutations, sequences from CHB participants were first compared with the OBI sequences for each respective subgenotype to identify potential OBI unique mutations and then compared to the GenBank reference sequences. To avoid polymorphisms that may represent subgenotype differences, sequences were compared at a subgenotype level [63,64]. Mutations that were unique to HBV isolated from OBI participants without appearing in any sequences from CHB patients or references from CHB were classified as occult-unique mutations [41,47,65–69].
3. Results

There were no statistically significant differences at baseline in terms of CD4+ T cell count, HIV viral load, liver enzymes, Fibrosis 4 (a noninvasive measure of liver scarring), or other clinical parameters between the CHB, OBI, and the HBV negative participants as reported in detail elsewhere [48,50]. The HBV viral loads were low in the OBI group with a median of 57.4 copies per mL versus 31,600 copies per mL in the CHB group as reported elsewhere [48]. HBV antigens results (HBsAg and HBeAg) and HBV antibodies (anti-HBc and anti-HBs) have been reported in detail elsewhere [48,50].

Of the 108 participants, HBV genomes from 50 (46.3%) individuals were successfully genotyped, including 25 of 36 (69.4%) from CHB and 25 of 72 (34.7%) from OBI. The low amplification rate likely reflects difficulties in amplifying longer fragments from individuals with low HBV levels such as occurs during OBI infections [64]. Twenty-seven samples were successfully genotyped from the whole 3.2 kb HBV genome (18 CHB and nine OBI) and 23 from the 3 kb fragment (seven CHB and 16 OBI). The circulating genotypes were 24 A1 (48%), 24 D3 (48%), and two E (4%), Supplementary Figure S1. For OBI participants, subgenotype A1 was found in 12 (48%) individuals, subgenotype D3 in 12 (48%), and genotype E in one (4%), with the same proportions found in CHB participants. OBI and CHB sequences clustered together (Figures 1 and 2). There were three serotypes that were found in this study; adw2 was found in all genotype As; ayw2 was found in all genotype Ds, except one, which was ay (could not be serotyped further); and ayw4 was found in the genotype E isolate.

Figure 1. A phylogenetic tree of the whole surface region (nucleotide (nt) 2854–835 from EcoRI site) of subgenotype A1 hepatitis B virus (HBV) sequences generated by Bayesian Evolutionary Analysis by Sampling Trees (BEAST). Strains from Botswana sequenced in the present study are shown in blue, while reference sequences are shown in black. Reference strains are designated by their accession number and country of origin, whereas Botswana sequences are designated by MA followed by a number and either CHB (for chronic HBV strains) or OBI (for occult HBV strains).
3.1. Results for Subgenotypes A1 and D3

Most sequences from this study grouped with the southern African A1, as shown in Figure 1. Only one sequence (MA94)—from an HBsAg positive participant—clustered with the Asian A1s. Mutations unique to MA94 were \textit{preS2}A7T, \textit{preS2}V17I, \textit{preS2}T31I (\textit{PreS2}), and \textit{preS1}A90V. There was no separate clustering of HBV sequences from HBsAg-negative and HBsAg-positive participants. Most sequences from Botswana had the subgenotype A1 unique, as in the \textit{preS1} region (Q54, V74, A86 and V91) and in the \textit{preS2} region (L32) \cite{70,71}, except for three sequences with mutations at position 86 (two had \textit{preS1}A86T and one was \textit{preS1}A86V). When comparing the sequences in \textit{preS1} ORF (positions five, six, and 25) and \textit{preS2} (position 48), all Botswana sequences had 5S except MA101 and MA87, which were 5P. At position six, all were 6A except MA94, MA95, and MA85, which were 6S, as well as MA89, which was 6I. At position 25, however, all were 25F, except two, MA95, and MA85, which were 25L. In the \textit{PreS2} at position 48, all study sequences had 48R except MA95 and MA85, which had 48T. Moreover, in the \textit{PreS2} region at position 38, the majority of Botswana sequences had 38T except eight (MA94, MA101, MA89, MA87, MA98, MA95, MA86, and MA85), which had 38I. Additionally, all study sequences clustering with Zimbabwean HBV sequences had 38T. The Botswana genotype D3s clustered with South African D3s, Figure 2.

![Figure 2](image-url)
3.2. Nucleotide Divergence

Nucleotide pairwise distances for each ORF were calculated using Molecular Evolutionary Genetics Analysis (MEGA) version 7.0.26, Table 2 [72]. There was a statistically significant difference \((p < 0.05)\) in the median nucleotide pairwise distances between CHB and OBI participants for most ORFs except in the preS2 (for both subgenotypes) and the X ORF (subgenotype D3), Table 2.

Table 2. Comparison of the median nnucleotide pairwise distance (%) between chronic and occult infections by open reading frame and subgenotype.

<table>
<thead>
<tr>
<th>ORF</th>
<th>CHB n = 12</th>
<th>OBI n = 12</th>
<th>p-Value</th>
<th>CHB n = 12</th>
<th>OBI n = 12</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pol</td>
<td>2.8 (1.8, 3.4)</td>
<td>0.5 (0.0, 1.4)</td>
<td>&lt;0.001</td>
<td>0.7 (0.4, 0.9)</td>
<td>0.4 (0.3, 0.5)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>S</td>
<td>1.2 (0.6, 2.1)</td>
<td>0.2 (0.1, 0.8)</td>
<td>&lt;0.001</td>
<td>0.3 (0.1, 0.9)</td>
<td>0.1 (0.0, 0.3)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>PreS1</td>
<td>3.2 (2.0, 5.5)</td>
<td>0.9 (0.0, 2.7)</td>
<td>&lt;0.001</td>
<td>1.1 (0.6, 1.1)</td>
<td>0.3 (0.1, 1.7)</td>
<td>0.001</td>
</tr>
<tr>
<td>PreS2</td>
<td>0.6 (0.0, 0.6)</td>
<td>0.0 (0.0, 1.2)</td>
<td>0.387</td>
<td>0.6 (0.0, 0.6)</td>
<td>0.0 (0.0, 1.2)</td>
<td>0.387</td>
</tr>
<tr>
<td>X</td>
<td>1.5 (1.1, 1.8)</td>
<td>11.8 (8.9, 23.3)</td>
<td>&lt;0.001</td>
<td>11.5 (1.9, 13.4)</td>
<td>8.6 (2.0, 11.5)</td>
<td>0.103</td>
</tr>
<tr>
<td>Core</td>
<td>4.0 (2.3, 5.8)</td>
<td>6.7 (2.8, 15.5)</td>
<td>&lt;0.001</td>
<td>0.9 (0.5–2.7)</td>
<td>1.1 (0.7, 1.7)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Abbreviations: CHB: Chronic hepatitis B; OBI: Occult hepatitis B infections; ORF: Open reading frame; Pol: Polymerase; S: Surface; PreS1: PreSurface 1; PreS2: PreSurface 2.

3.3. HBV Escape Mutations in the S ORF of OBI Versus CHB Sequences

Ten escape mutations associated with vaccine, diagnostic, or immunoglobulin therapy failure were identified in the S ORF. These include four mutations in OBI sequences (\(\sigma_K122R, \sigma_C124Y, \sigma_Q129H, \) and \(\sigma_K160N\)) and seven in CHB sequences (\(\sigma_P120S, \sigma_M103I, \sigma_Q129R, \sigma_G130N, \sigma_T140I, \sigma_G145R, \) and \(\sigma_K160N\)). One escape mutation—\(\sigma_K160N\)—occurred in both groups. The frequency of all mutations was one of 25 (4%) sequences, Table 3.

Table 3. HBV escape mutations in the S ORF of OBI versus CHB sequences.

<table>
<thead>
<tr>
<th>Mutation</th>
<th>HBV Type</th>
<th>Reported Impact</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\sigma_P120S)</td>
<td>OBI</td>
<td>Vaccine, detection escape</td>
<td>[73]</td>
</tr>
<tr>
<td>(\sigma_K122R)</td>
<td>OBI</td>
<td>Decreased HBsAg expression, diagnostic escape</td>
<td>[42,74]</td>
</tr>
<tr>
<td>(\sigma_C124Y)</td>
<td>OBI</td>
<td>Immunoglobulin therapy escape, decreases antigenicity and viral secretion</td>
<td>[73,75]</td>
</tr>
<tr>
<td>(\sigma_Q129R)</td>
<td>OBI</td>
<td>Vaccine and immunoglobulin therapy escape, decreases viral secretion</td>
<td>[73,76]</td>
</tr>
<tr>
<td>(\sigma_K160N)</td>
<td>OBI/CHB</td>
<td>Decreases antigenicity and viral secretion</td>
<td>[77]</td>
</tr>
<tr>
<td>(\sigma_M103I)</td>
<td>CHB</td>
<td>Decreases antigenicity</td>
<td>[42]</td>
</tr>
<tr>
<td>(\sigma_Q129R)</td>
<td>CHB</td>
<td>Vaccine and detection escape and decrease viral secretion</td>
<td>[73,75]</td>
</tr>
<tr>
<td>(\sigma_G130N)</td>
<td>CHB</td>
<td>Diagnostic and vaccine escape</td>
<td>[78,79]</td>
</tr>
<tr>
<td>(\sigma_T140I)</td>
<td>CHB</td>
<td>Decreases viral secretion</td>
<td>[75]</td>
</tr>
<tr>
<td>(\sigma_G145R)</td>
<td>CHB</td>
<td>Vaccine, detection and immunoglobulin therapy escape, decreased antigenicity and viral secretion</td>
<td>[75]</td>
</tr>
</tbody>
</table>

Abbreviations: CHB: Chronic hepatitis B; OBI: Occult hepatitis B infections; S: Surface.

3.4. Occult-Unique Mutations

Occult-unique mutations were those mutations that only appeared in HBV sequences from OBI participants and not in CHB or reference sequences. For subgenotype A1, 12 HBV strains from OBI positive participants were compared with 119 CHB strains (12 CHB from this study + 107 GenBank CHB references). For subgenotype D3, 12 OBI strains were compared with 97 CHB strains (12 from this study + 85 from GenBank reference sequences). A total of 43 OBI unique mutations were identified in this study. Several of these mutations were reported previously. However, 26 (60.5%) were novel. Each mutation was found in a single strain, Table 4. The OBI unique mutations were found in HBV strains from 14 participants most of whom (12) harbored multiple mutations, Table 5.
Table 4. OBI unique mutations in the different ORFs.

<table>
<thead>
<tr>
<th>ORF/Region</th>
<th>Variant</th>
<th>Subgenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>PreS1</td>
<td>preS1 S78N</td>
<td>D3</td>
</tr>
<tr>
<td>PreS2</td>
<td>preS2 F22P, preS2 F22H</td>
<td>D3</td>
</tr>
<tr>
<td>S</td>
<td>L97P, T114I, C124Y, N131K, P217L</td>
<td>A1</td>
</tr>
<tr>
<td>S</td>
<td>Q129H</td>
<td>D3</td>
</tr>
<tr>
<td>Pol</td>
<td>N120Y, K155R, W64R, P103S, T128I, Y257F</td>
<td>A1</td>
</tr>
<tr>
<td>X</td>
<td>P11S, S31A, Q87L, S101P, L116V</td>
<td>A1</td>
</tr>
</tbody>
</table>

Abbreviations: ORF: Open reading frame; S: surface; preS1: Pre-Surface 1; preS2: Pre-Surface 2; Pol: Polymerase; TP: Terminal protein; SP: spacer; RT: Reverse transcriptase; RH: RNase H; C: core. Novel OBI unique mutations are shown in bold.

Table 5. OBI unique mutations in different participants.

<table>
<thead>
<tr>
<th>PID</th>
<th>Subgenotype</th>
<th>Variant</th>
</tr>
</thead>
<tbody>
<tr>
<td>MA121</td>
<td>D3</td>
<td>preS2 F22H</td>
</tr>
<tr>
<td>MA122</td>
<td>D3</td>
<td>preS1 S78N, preS2 F22P, W64R, P103S</td>
</tr>
<tr>
<td>MA123</td>
<td>D3</td>
<td>D2A, E64K</td>
</tr>
<tr>
<td>MA127</td>
<td>D3</td>
<td>T128I, W102Y, P103V</td>
</tr>
<tr>
<td>MA125</td>
<td>D3</td>
<td>I3R, D4Y</td>
</tr>
<tr>
<td>MA107</td>
<td>A1</td>
<td>N131K, L140I</td>
</tr>
<tr>
<td>MA102</td>
<td>A1</td>
<td>T225A</td>
</tr>
<tr>
<td>MA99</td>
<td>A1</td>
<td>E117Stop, R127H</td>
</tr>
<tr>
<td>MA126</td>
<td>D3</td>
<td>Q129H, Y257F, Q87L</td>
</tr>
<tr>
<td>MA100</td>
<td>A1</td>
<td>I81M, S11A</td>
</tr>
<tr>
<td>MA101</td>
<td>A1</td>
<td>T114I, A329T</td>
</tr>
</tbody>
</table>

Abbreviations: S: surface; preS1: Pre-Surface 1; preS2: Pre-Surface 2; Pol: Polymerase; TP: Terminal protein; SP: spacer; RT: Reverse transcriptase; RH: RNase H; C: core.

3.5. Signature Amino Acid Analysis

The VESPA available at https://www.hiv.lanl.gov/content/sequence/VESPA/vespa.html was used to determine signature aa occurring in HBV sequences from OBI participants compared to CHB sequences. CHB sequences from the current study were used as background, and all analyses were performed by subgenotype. A total of 16 signature aa were detected, 12 in subgenotype A1 and four in subgenotype D3, Table 6. The PreS2 and C ORFs had no signature aa in either subgenotype.

3.6. Immune Selection Pressure

There were 26 negatively selected and no positively selected codons in the C ORF. Three of the negatively selected codons in the CHB group (four, 97, and 45) were in positions with OBI-unique mutations in the respective subgenotype. Two positions were under negative selection in both CHB and OBI participants (96, 122). None of the signature aa overlapped with codons under negative selection pressure. Five of the codons (four, seven, 68, 69, and 122) were within known HBV T cell epitopes: 1–20, 50–69, and 121–140. In the Pol ORF, there were 50 negatively selected codons and one positively selected codon. Only one OBI-associated mutation (r81) overlapped with a codon under negative selection in the CHB participants, and two codons were negatively selected in both CHB and OBI participants. In the PreS1 and S region, two and one codons, respectively, were negatively selected in both CHB and OBI participants. In the Pre S1 region, seven of the codons were in HBV immune epitope positions: (28, 29) in the 21–30 T cell epitope (42, 44, 46) in the 29–48 T cell epitope and (85
and 90) in the 81–95 T cell epitope. In the X region, two codons were in epitopes positions (86 and 95): B cell epitope 85–110, whereas in the S region, one codon (148) was situated in a B cell epitope position: 122–148, Supplementary Table S1. There were significantly more negatively selected codons in the CHB than in the OBI (p = 0.031) sequences.

<table>
<thead>
<tr>
<th>ORF</th>
<th>Sub Genotype</th>
<th>Signature Amino Acid</th>
<th>Number in OBI Participants *</th>
<th>Number in CHB Participants *</th>
</tr>
</thead>
<tbody>
<tr>
<td>PreS1</td>
<td>D3</td>
<td>preS1P65L</td>
<td>11</td>
<td>5</td>
</tr>
<tr>
<td>PreS1</td>
<td>A1</td>
<td>PreS1P94T</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>Surface</td>
<td>A1</td>
<td>S113H</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>Surface</td>
<td>A1</td>
<td>S1194V</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>Pol</td>
<td>D3</td>
<td>rH54V</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Pol</td>
<td>A1</td>
<td>tH87Q</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>Pol</td>
<td>A1</td>
<td>sP11A</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>Pol</td>
<td>A1</td>
<td>sA68P</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>Pol</td>
<td>A1</td>
<td>sP97H</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>Pol</td>
<td>A1</td>
<td>rI103V</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>Pol</td>
<td>A1</td>
<td>rQ139H</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>Pol</td>
<td>A1</td>
<td>rS32N</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>Pol</td>
<td>A1</td>
<td>rK333Q</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>Pol</td>
<td>A1</td>
<td>rP2S</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>X</td>
<td>D3</td>
<td>xG27R</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>X</td>
<td>D3</td>
<td>xF30L</td>
<td>7</td>
<td>3</td>
</tr>
</tbody>
</table>

Abbreviations: ORF: Open reading frame; OBI: occult hepatitis B infections; CHB: chronic hepatitis B infections; Pol: Polymerase; tp: Terminal protein; sp: Spacer; rt: Reverse transcriptase; rh: RNase H; S: Surface * is n = 12.

4. Discussion

This is the first study to report whole genome and nearly whole genome HBV sequences from Botswana and their molecular characterization. We also report the HBV genotypes in participants with OBI, and we identified mutations associated with OBI.

Subgenotypes A1, D3, and genotype E were identified in both OBI and CHB participants, which were also isolated from CHB participants in previous studies carried out in Botswana [4,80]. Most HBV sequences from Botswana clustered with other African sequences both at whole genome and at whole surface region level, as indicated in a study by Makondo et al. [81]. Similar to the aforementioned South African study, in which some isolates clustered with Asian sequences, one sequence also clustered with Asian sequences [81]. In further concordance with the South African study, sequences from HBsAg-positive and HBsAg-negative participants clustered together [81].

There are several mechanisms that have been implicated in OBI, including mutations in the HBV genome. However, there are few studies that have undertaken mutational analysis on nearly whole genomes of OBI strains owing to the difficulty in amplifying OBI samples with low viral loads [35,36,64,82–84]. Hence, most studies have focused on fragments of the HBV genome, especially the surface gene [41,42,64,85–87]. This is the first study to perform mutational analysis of nearly whole HBV genome sequences from Botswana.

The S ORF is evaluated most frequently for identification of OBI associated mutations. Some mutations in this ORF may lead to decreased detection by enzyme-linked immunosorbent assay (ELISA) kits, whereas some mutations demonstrated decreased secretion of HBsAg or increased retention of HBsAg [41,42,64,85–87]. All OBI associated mutations identified in the S ORF in this study (Table 4) have been reported previously [88–92]; however, the S region has been extensively studied [41,42,64,85–87]. The sL97P and sT114I mutations are located in an immunodominant region, the major hydrophilic region (position 99 to 169) [93]. Furthermore, the sN131K and sP217L mutations have been associated with diagnostic escape [89,94]. One of the mutations (sC124Y) has been reported
and characterized in OBI sequences from blood donors in China and has been found to diminish both viral antigenicity and viral secretion [75]. There are multiple mutations that have been reported before in the same position and found to impact OBI phenotype in both a similar or different manner, e.g., (sC124R, sC124Y, and sC124A); sC124A was found to decrease extracellular HBsAg by decreasing viral secretion, whereas sC124R and sC124Y decreased antigenicity and viral secretion [75,76]. Cysteine forms disulfide bridges, which are critical in the formation of the HBsAg structure, hence, loss of cysteine might alter the structure and impact immunogenicity [95].

Conversely, for D3 isolates, there was only one OBI-associated mutation (sQ129H), which has been reported and functionally characterized before and found to decrease viral secretion [76]. Other studies have, however, reported OBI associated mutations, which have never been reported before, such as a study in South Africa [41]. However, there are differences between this study and the current study. In the South African study, most OBI mutations were from subgenotype A2, a subgenotype not present in the current study [41]. Another study in South Africa also identified three new OBI-associated mutations for genotype A participants (sY72H, sL82T, and sA128T), which were not observed in the current study [65].

In Botswana, there were no OBI-unique mutations in the PreS1 ORF for subgenotype A1 participants. The absence of OBI associated mutations has been reported before; for example, a study in Turkey found no OBI associated mutations in the S gene of female sex workers [96]. Some studies have reported deletions, which reduced HBsAg production in OBI participants, whereas other studies reported point mutations in this ORF [41,44,65,97]. The studies that reported point mutations concur with the present study, which reported one OBI-associated mutation in PreS1 for subgenotype D3. The preS1S78N mutation was novel. A different variant (preS1S78G) was observed at the same position in subgenotype A1 participants from South Africa [41].

There were two OBI-unique mutations in the PreS2 ORF of the subgenotype D3 sequences, as in other studies, which also found point mutations in this ORF [41,65,96,98]. A different variant (preS2F122L) was reported at the position of the two OBI associated mutations and has been linked with HCC [99,100]. The OBI-unique mutations in the current study are within the attachment region for human albumin (aa 17–28). Therefore, they may reduce HBV infectivity, since preS2 is responsible for mediating HBV attachment to hepatocytes [101]. Some studies documented deletions in the preS2 ORF, which may reduce HBsAg secretion [98,102]. The two mutations preS2F22P and preS2F22H were novel. The loss of phenylalanine, an aromatic amino acid, might impact protein structure, as aromatic aa play a role in protein structure stabilization [103].

The OBI associated mutations have also been described in the C ORF [22]. In the current study, there were 16 OBI-unique mutations in this ORF, and some of these mutations were located in functionally relevant regions. Three of these mutations (cD2A, cI3R, and cD4Y) spanned CD4+ T cell epitope (aa 1–20), and two other mutations (cE64K and cI59L) from this study were located within another CD4+ T cell epitope (aa 50–69) [104,105]. Furthermore, mutations cE117Stop and cI59L were also found in a CD4+ T cell epitope (aa 117–131) [104,105]. There is a mutation (cR127H), which was located in the CD4+ T cell epitope (aa 120–140) [104,105]. Some of the 16 OBI mutations in the core region have been reported before. For instance, cE64K has been reported in OBI-positive children born to HBsAg positive mothers in Iran [22], whereas cE46D was found in OBI-positive Chinese patients [35]. The aforementioned study [35] also reported different variants at the same residue as some of the OBI associated mutations (aa positions 26, 59, 74, 87, and 113) [35]. cV74N, cS87N, and cF97I have also been reported previously [106], with cF97I also detected in HBV from India [107]. Mutation cR127H has been shown to block capsid formation [108]. In the present study, there were 10 novel OBI-unique mutations in the core region: cS26P, cD32G, cP45S, cI59L, cD2A, cI13R, cD4Y, cW102V, cF103V, and cE117Stop. Position 102 plays a role in capsid assembly [109].

The X ORF is essential for transcription, hence, mutations in this ORF may affect viral replication [64]. In this study, there were seven OBI associated mutations identified in the X region, four of which have been previously reported. Mutation xS31A has been reported previously,
and phosphorylation of the X protein occurs at serine 31 \cite{110,111}. Mutations xQ87L, xS101P, and xL116V have also been reported in the literature \cite{99,112}. In the present study, three novel OBI-unique mutations were detected in this ORF (xS11A, xV15I, and xP11S). Other studies in India reported OBI-associated mutations in the X ORF \cite{113}. Some studies have reported deletions in the X region, which are linked to OBI. For instance, a Korean study documented an 8 bp deletion in OBI strains \cite{36}. These deletions were shown to decrease HBsAg and HBV secretion and therefore may be responsible for the OBI phenotype \cite{36}. In spite of this, in the present study, similar to other studies, there were no long stretches of deletions in the OBI strains \cite{22}.

Several studies have reported mutations in the Pol gene, which lead to drug resistance, and because of the overlapping nature, the mutations also simultaneously lead to changes in the S region, which may affect HBsAg production \cite{114}. For example, studies in India and South Africa documented drug resistance mutations in OBI-positive participants \cite{113,115}. In this study, rtT128I has also been reported similar to several studies including in chronic participants, and it has been linked to lamivudine drug resistance \cite{116--121}. One of the mutations found in this study (spL140I) has been reported in HIV-infected participants in Germany but has not been characterized yet \cite{89}. A study in China reported a 218 bp deletion in the Pol region of OBI positive adults, which might explain the low level of HBV viral load in the OBI participants \cite{67}. In the current study, there were no deletions. On the other hand, the Chinese study found no OBI-associated mutations in the Pol region \cite{44}. There were 10 novel OBI mutations in this region: tpN120Y, tpK155R, spW64R, spS91T, spP103S, spS133G, rtT225A, rtY257F, rtA329T, and rhI81M.

In most ORFs, more codons were under negative immune selection pressure in the CHB participants compared to the OBI participants. These results concur with a South African study, which found more negatively selected codons in the surface and Pol region in chronic participants compared to their OBI counterparts \cite{41}. The reduced negative selection in the OBI participants might be due to the low HBV DNA, a characteristic of OBI \cite{19,41}. Several of the negatively selected codons identified in this study are similar to those described in the South African study \cite{41}. The overlap might be due to some similarities between the populations owing to the close proximity of the two countries. In contrast to the 16 negatively selected codons, which occurred in sites with OBI-unique mutations, in the current study only four were detected. These differences may be because the former study had more CHB participants (31 vs. 24) and also variations in HBV genotypes studied (A1 and A2 versus A1 and D3) \cite{41,122}. Generally, there were also more negatively selected sites in genotype A compared to genotype D. Differences between genotypes even in mutations have been reported before \cite{122}. Some of the codons under negative immune selection pressure were in known HBV epitopes \cite{123--127}. Furthermore, nucleotide divergence was higher in the CHB group compared to the OBI group for most ORFs, Table 2, which concurs with other studies \cite{113}. These differences in diversity may be due to immune pressure in the CHB group, as HBsAg triggers a greater immune response, thereby increasing the number of mutations arising, as the virus responds to evade the immune system. The low diversity seen in the OBI group may be due to the lack of HBsAg.

Several limitations of this study should be noted. The modest sample size may not reflect all aspects of HBV sequences circulating in Botswana or Southern Africa. Furthermore, the preC region was not analyzed, because it was not amplified in most participants. As most study participants were HIV-positive, findings from this study may not be generalizable to HBV mono-infected participants. Future HBV full-length studies on a larger sample size, including HBV mono-infected individuals, are warranted.

In conclusion, mutations in HBV from OBI participants in Botswana have been determined for the first time. Some of the OBI-unique mutations reported here have been functionally characterized before and found to impact HBsAg secretion/production and HBV virion secretion, and hence have an impact in the OBI phenotype. Other OBI mutations have been reported before, but their functional relevance is still unknown. In addition to these, there were also mutations that were found in this study that have not been reported previously. This study, therefore, has added to the body of knowledge
of OBI-associated mutations by reporting novel OBI-unique mutations. Knowledge of OBI-unique mutations is important considering that with increased knowledge of OBI, diagnostic and preventative measures may be put in place in order to eliminate HBV per WHO goals.

Supplementary Materials: The following are available online at http://www.mdpi.com/2073-4425/9/9/453/s1, Table S1: Immune selection pressure results determined using fixed effects likelihood in the DataMonkey Figure S1: A BEAST tree of full length HBV sequences from Botswana and reference sequences from GenBank.


Funding: This work was supported by the Southern Africa Consortium for Research Excellence (Grant number 087537/F/08/Z) and National Institutes of Health (NIH) Fogarty International Center (Grant number 5D43TW009610). The research was also partly funded by the Sub-Saharan African Network for TB/HIV Research Excellence (SANTHE), a DELTAS Africa Initiative grant number DEL-15-006. The DELTAS Africa Initiative is an independent funding scheme of the African Academy of Sciences (AAS)’s Alliance for Accelerating Excellence in Science in Africa (AESA) and is supported by the New Partnership for Africa’s Development Planning and Coordinating Agency (NEPAD Agency), with funding from the Wellcome Trust grant number 107752/Z/15/Z and the UK government. The views expressed in this publication are those of the author(s) and not necessarily those of the AAS, NEPAD Agency, Wellcome Trust, or the UK government. S.M. was supported in part by the Fogarty International Center and National Institutes of Mental Health of the National Institutes of Health under Award Number D43 TW010543. AK was funded by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG). T.G.B received funding from the DFG and the University Research Committee (URC). The funders had no role in the study design, data collection, and decision to publish, or in the preparation of the manuscript.

Acknowledgments: We acknowledge Bomolemo and Tshipidi study participants.

Conflicts of Interest: The authors declare no conflict of interest.

References


42. Martin, C.M.; Welge, J.A.; Rouster, S.D.; Shata, M.T.; Sherman, K.E.; Blackard, J.T. Mutations associated with occult hepatitis B virus infection result in decreased surface antigen expression in vitro. *J. Viral Hepat.* 2012, 19, 716–723. [CrossRef] [PubMed]


60. Los Alamos National Laboratory, VESPA. Available online: https://www.hiv.lanl.gov/content/sequence/ VESPA/vespa.html (accessed on 30 May 2017).


62. Bell, T.G.; Kramvis, A. Bioinformatics tools for small genomes, such as hepatitis B virus. *Viruses* 2015, **7**, 781–797. [CrossRef] [PubMed]


65. Martin, C.M.; Welge, J.A.; Shire, N.J.; Rouster, S.D.; Shata, M.T.; Sherman, K.E.; Blackard, J.T. Genomic variability associated with the presence of occult hepatitis B virus in HIV co-infected individuals. *J. Viral Hepat.* 2010, **17**, 588–597. [CrossRef] [PubMed]


67. Chen, S.J.; Zhao, YX; Fang, Y.; Xu, W.Z.; Ma, YX; Song, ZW; Teng, X.; Gu, H.X. Viral deletions among healthy young Chinese adults with occult hepatitis B virus infection. *Virus Res.* 2012, **163**, 197–201. [CrossRef] [PubMed]


81. Makondo, E.; Bell, T.G.; Kramvis, A. Genotyping and molecular characterization of hepatitis B virus from human immunodeficiency virus-infected individuals in southern Africa. *PLoS ONE* 2012, 7. [CrossRef] [PubMed]


111. Khattar, E.; Mukherji, A.; Kumar, V. Akt augments the oncogenic potential of the HBx protein of hepatitis B virus by phosphorylation. *FEBS J.* 2012, 279, 1220–1230. [CrossRef] [PubMed]


© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).