Hepatitis B virus infection in post-vaccination South Africa:
ocult HBV infection and circulating surface gene variants

Edina Amponsah-Dacosta¹, Ramokone L. Lebelo¹, J. Nare Rakgole¹, Selokela G. Selabe¹, Maemu P. Gededzha¹, Simnikiwe H. Mayaphi², Eleanor A Powell³, Jason T. Blackard³, and M. Jeffrey Mphahlele¹#

¹HIV and Hepatitis Research Unit, Department of Virology, University of Limpopo (Medunsa Campus) and National Health Laboratory Service, Pretoria, South Africa
²Department of Medical Virology, University of Pretoria/Tshwane Academic Division of NHLS, Pretoria, South Africa
³Division of Digestive Diseases, Department of Internal Medicine, University of Cincinnati College of Medicine, Cincinnati, Ohio, USA

# Correspondence:

MJ Mphahlele
Department of Virology, University of Limpopo (Medunsa Campus)
P.O. Box 173, MEDUNSA, 0204
Pretoria, South Africa
Tel +27 12 521 4227 or 4569 Fax: +27 12 521 5794
Email: jeffrey.mphahlele@ul.ac.za or jeffmphahlele@yahoo.com

INTRODUCTION

South Africa is hyper-endemic for hepatitis B virus (HBV) infection which is associated with fatal chronic sequelae such as liver cirrhosis, end-stage liver disease and hepatocellular carcinoma (HCC) (Kramvis and Kew, 2007; McMahon, 2009).
Prior to hepatitis B vaccine introduction, the burden of hepatitis B in the country was chiefly driven by the high prevalence (>10%) of childhood (<5 years) infections accounting for ~3 million potentially fatal chronic HBV infections (Vardas et al, 1999; Kew, 2008).

In keeping with the World Health Assembly’s 1992 recommendation for controlling the hepatitis B burden in hyper-endemic regions, the hepatitis B vaccine was introduced into the South African Expanded Programme on Immunization (EPI-SA) in April 1995 (WHO, 1992, South African National Department of Health, 1995). Subsequent studies have reported a decline in the prevalence of chronic HBV infection from 10% to <0.5% in vaccinated cohorts (Tsebe et al, 2001; Schoub et al, 2002; Simani et al, 2008). What threatens the benefits of vaccination is the emergence HBV S gene variants due to immune selection pressure from the HBV vaccines (Zuckerman and Zuckerman, 2003). These variants, which may exhibit immune, diagnostic and vaccine escape, can also be transmitted (Carman, 1997; Echevarría and Avellón, 2006; Yokosuka and Arai, 2006). While this is a very rare event, vaccine escape mutations have been detected in several countries with longstanding hepatitis B vaccination programmes, such as Taiwan (Hsu et al, 1999; Chang, 2010) and The Gambia (Fortuin et al, 1994; Whittle et al, 1995).

It is reasonable to expect the remarkable impact the hepatitis B vaccines had on HBV prevalence to be similar on occult HBV infection (OBI) which is also prevalent in South Africa, especially in the human immunodeficiency virus (HIV) infected population (Mphahlele et al, 2006; Lukhwareni et al, 2009). The Taormina experts meeting defines OBI as the presence of HBV DNA in the liver with detectable
(<200IU/mL) or undetectable HBV DNA in the serum of individuals testing negative for HBV surface antigen (HBsAg) by current available assays (Raimondo et al, 2008). Those infections that fit this criteria but have a detectable HBV DNA of >200 IU/mL are referred to as false OBI (Reference). Clinically, OBI may be benign but is also associated with increased risk of reactivation of HBV infection in immunocompromised individuals (for example during HIV infection) and can hasten progression of liver disease and hepatocarcinogenesis (Chemin and Trèpo, 2005; Romero et al, 2011). The public health concern is in the risk of transmission of HBV from asymptomatic and misdiagnosed occult HBV infected individuals to susceptible persons (Ocana et al., 2011).

The aim of this study was to examine the HBV S gene variants and occult HBV infections circulating in a South African population post-vaccine introduction.

**MATERIALS AND METHODS**

From a previous hepatitis B serosurvey, 201 of 1206 serologically exposed (HBsAg and/or anti-HBc positive) serum samples were identified for the purpose of this study. Patient records were obtained through the National Health Laboratory Services (NHLS) Virology Diagnostic Unit, Dr. George Mukhari Academic Hospital. The serum samples were stratified by age with reference to the year (1995) the hepatitis B vaccine was introduced into EPI-SA, into a post-vaccine introduction population (POVP) and a pre-vaccine introduction population (PRVP). Where HIV records were available, a subset analysis of results was performed. The hepatitis B
vaccination status of patients was not collected for this study so as to mimic population-based nationwide hepatitis B serosurveys.

Ethical clearance to conduct this study was approved by the Medunsa Research and Ethics Committee (MREC) (MREC/P/22/2012: PG). In order to ensure patient confidentiality, all personal identifiers were removed from the research database.

**HBV DNA testing**

All 201 serum samples were tested for HBV DNA using a quantitative real-time PCR (qPCR) assay performed on the LightCycler® 2.0 Real-Time PCR System (Roche, Germany). The 5X HOT FIREPol® EvaGreen® qPCR Mix Plus (no ROX) (Solis BioDyne, Estonia) reagent was used together with primers designed to target the HBV basal core promoter (BCP) region (BcP1: 5’-ACCACCAAATGCCCTAT-3’ and BcP2: 5’-TTCTGCGACGCGGCGA-3’) (Ho et al, 2003). Prior to testing, a standard curve was generated using duplicate 10 fold serial dilutions of ACCURUN® 325 HBV DNA positive control (SeraCare Life Sciences, USA) with an initial viral load of 3.04 x 10⁶ IU/ml. Using these dilutions the lower limit of detection of the assay was 3.04 x 10⁰ IU/ml. A conversion factor of 2.8 IU/ml to copies/ml was used for all viral loads determined (Paraskevis et al, 2010).
**HBV S gene amplification and sequencing**

For amplification of HBV DNA in the S region, a conventional nested PCR assay was performed using two sets of primer pairs (first round: S1 5’-CCTGCTGGTGCTCCAGTTCC-3’, S2Na 5’-CCACCATTCK(G/T)TTGACATACTTTCCA-3’ and second round: S6Bs 5’-GATCCGAGGACTGGGGAC-3’, S7Ps 5’-GGTTAGGGTTAAATGTATAC-3’) in a two-step reaction for all HBV DNA positive samples with viral loads >35 IU/mL (detection limit by this PCR assay). The cycling conditions used for amplification of the S gene were as outlined previously by Mphahlele et al (2006). Confirmation of the desired amplicon (~681 bp) was performed by gel electrophoresis on a 2% agarose gel and then direct sequencing of the gene carried out using the SpectruMedix SCE 2410 Genetic Analysis System (SpectruMedix LLC, PA).

**Phylogenetic analysis of HBV S gene sequences**

All sequences were edited and aligned using ChromasPro 1.49Beta (Technelysium Pty. Ltd.) and BioEdit Sequence Alignment Editor (Hall, 1999) respectively, and then analysed for both nucleotide and amino acid variations. For phylogenetic analysis, representative sequences from different HBV sub-genotypes were retrieved from the GenBank database for comparison with study sequences. All alignments were performed using the neighbor-joining method implemented in Clustal X (Thompson et al, 1997). The statistical robustness and reliability of the branching order was confirmed by bootstrap analysis using 1000 replicates (Felsenstein, 1985). Phylogenetic inference was also performed using a Bayesian Markov chain Monte
Carlo (MCMC) approach as implemented in the BEAST v1.8.0 program (Drummond et al, 2012) under an uncorrelated log-normal relaxed molecular clock using the general time-reversible model with nucleotide site heterogeneity estimated using a gamma distribution. The MCMC analysis was run for a chain length of 200,000,000 with sampling every 5,000th generation. Results were visualized in Tracer v1.5 to confirm chain convergence, and the effective sample size (ESS) was calculated for each parameter. All ESS values were >500 indicating sufficient sampling. The maximum clade credibility tree was selected from the posterior tree distribution after a 10% burn-in using TreeAnnotator v1.8.0.

RESULTS

Study population

Of the 201 enrolled participants, 62 (..%) belonged to POVP and 139 (..%) belonged to PRVP. Of the total population, 69.7% (140/201) were females, while records of patients’ gender were unavailable for 1.0% (2/201). The mean ages were 9.2 years (SD=4.9) and 22.7 years (SD=2.2) for the POVP and PRVP, respectively (Table 1). Serum samples were collected from health facilities in five of the nine provinces in South Africa, including Gauteng (49.8% [100/201]), North West (40.3% [81/201]), Mpumalanga (6.0% [12/201]), Limpopo (3.5% [7/201]), and the Northern Cape (0.5% [1/201]).
**Prevalence of OBI**

The prevalence of OBI was evaluated in all HBsAg negative but seropositive (i.e. positive for anti-HBc with or without anti-HBs) samples. Overall, 85.5% (53/62) of the POVP and 82.7% (115/139) of the PRVP fit this serological profile. Of those, HBV DNA was detectable in 66.0% (35/53) of the POVP and 70.4% (81/115) of the PRVP with 40.0% (14/35) and 37.0% (30/81) respectively having the anti-HBc only serologic profile while the rest tested positive for both anti-HBc and anti-HBs. Based on detection of serum viral loads <200 IU/mL, the prevalence of true OBI was 54.7% (29/53) in the POVP and 68.7% (79/115) in the PRVP (Table 1). Of these, 44.8% (13/29) and 35.4% (28/79) of the POVP and PRVP respectively had the anti-HBc only serologic profile.

In the HIV-infected subset, 92.3% (12/13) of the POVP and 91.7% (22/24) of the PRVP tested HBsAg negative but seropositive for anti-HBc with or without anti-HBs and of those, HBV DNA was detectable in 91.7% (11/12) and 63.6% (14/22) of the POVP and PRVP, respectively. All of those in the PRVP (63.6%) were true occult cases, while 75.0% (9/12) of the POVP had viral loads <200 IU/mL. The anti-HBc only serologic profile was evident in 55.6% (5/9) and 50.0% (7/14) of the POVP and PRVP respectively. In the HIV uninfected subset, all samples from the POVP (10/10) and 90.0% (9/10) of the PRVP were HBsAg negative but seropositive, with HBV DNA detected in 70.0% (7/10) and 88.9% (8/9) of the POVP and PRVP, respectively. The prevalence of true OBI in the HIV uninfected subset was 50.0% (5/10) in the POVP and 88.9% (8/9) in the PRVP, with the anti-HBc only serologic profile detected in 20.0% (2/10) and 22.2% (2/9) respectively.
Table 1: Demographics and HBV infections in the POVP versus the PRVP

<table>
<thead>
<tr>
<th></th>
<th>POVP (n=62)</th>
<th>PRVP (n=139)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Average age (range)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HBsAg positive</td>
<td>14.5% (9/62)</td>
<td>17.3% (24/139)</td>
<td>0.63</td>
</tr>
<tr>
<td>HBsAg negative*</td>
<td>85.5% (53/62)</td>
<td>82.7% (115/139)</td>
<td>0.63</td>
</tr>
<tr>
<td><em><em>HBsAg negative</em> and HBV DNA positive</em>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anti-HBc only with positive HBV DNA</td>
<td>66.0% (35/53)</td>
<td>70.4% (81/115)</td>
<td>0.57</td>
</tr>
</tbody>
</table>

**True Occult**

(HBV DNA <200 IU/mL)

- Anti-HBc only with true occult
  - 54.7% (29/53)
  - 68.7% (79/115)
  - 0.08

False occult

(HBV DNA >200 IU/mL)

HIV status

- Anti-HBc only serology
  - True OBI
  - False OBI

*Positive for anti-HBc with or without anti-HBs
**Phylogenetic analysis of the HBV S gene**

A total of 47 of 149 HBV DNA positive samples were selected for PCR amplification of HBV DNA in the S region based on viral loads >35 IU/ml (detection limit by the PCR assay). The HBV S gene was successfully amplified and sequenced for 37/47 isolates, 9 of these in the POVP and 28 in the PRVP. Phylogenetic analysis showed that the majority of these isolates including all 9 POVP isolates were subgenotype A1 with just two other isolates (PRE-1826 and PRE-8985) forming a monophyletic cluster with the subgenotype A2 reference data set, supported by a posterior probability value of 1.0 (Figure 1). A subgenotype D4 isolate was also identified from the PRVP, clustering with reference subgenotype D4 isolates from Haiti and Australia (data not shown).

**Molecular analysis of variations within the HBV S gene**

A total of 20 different amino acid variations were observed, most notably F8L (PRE-9625) and M103I (POST-0225, POST-4028, PRE-1082 and POST-4288). The following variations were unique to HBV isolates from “false” OBI (i.e. those that tested HBsAg negative but had high viral loads [1081.48 – 10962.96 IU/mL] comparable to that in chronic HBV infections): S45P (PRE-3106), P70H (POST-4162) and V168A+P217L (PRE-5603). Two nonsense mutations were also observed to the end of the S protein of isolates which tested HBsAg positive: Y200* (isolate POST-5162) and L206* (PRE-1082), caused by a stop codon (TAA) inserted by T600A and T647A nucleotide substitutions in the respective isolates.
Fig 1: Bayesian inference tree created in BEAST showing the phylogenetic analysis of HBV subgenotype A1 and A2 sequences. Study sequences are shown in red (PRVP) and blue (POVP) while reference sequences can be identified by their accession numbers as available in GenBank. Posterior probabilities are indicated at the nodes of the tree.
The antigenic region of the HBsAg – the Major Hydrophilic Region (MHR, aa110-aa165) containing the viral “α” epitope (aa124-aa147) – was also investigated for amino acid variations. Table 2 shows the variations observed, including 3 neutral substitutions: a K122R (PRE-4065), N131T (PRE-6964), and T143S (PRE-1177). Within the MHR but outside the “α” epitope, an A159P amino acid substitution was observed for isolate POST-1474 and an E164D for isolate PRE-7532.

<table>
<thead>
<tr>
<th>Isolate</th>
<th>HBV subgenotype</th>
<th>Wild type amino acids</th>
<th>Amino acid position</th>
<th>Variant amino acid</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRE-4065</td>
<td>A1</td>
<td>Lysine/Arginine (K/R)</td>
<td>122</td>
<td>Arginine (R)</td>
</tr>
<tr>
<td>PRE-6964</td>
<td>D4</td>
<td>Asparagine/Threonine (N/T)</td>
<td>131</td>
<td>Threonine (T)</td>
</tr>
<tr>
<td>PRE-1177</td>
<td>A1</td>
<td>Threonine/Serine (T/S)</td>
<td>143</td>
<td>Serine (S)</td>
</tr>
<tr>
<td>POST-1474</td>
<td>A1</td>
<td>Alanine/Glycine (A/G)</td>
<td>159</td>
<td>Proline (P)</td>
</tr>
<tr>
<td>PRE-7532</td>
<td>A1</td>
<td>Glutamic acid (E)</td>
<td>164</td>
<td>Aspartic acid (D)</td>
</tr>
</tbody>
</table>

**DISCUSSION**

The prevalence of true OBI in the South African population following introduction of universal hepatitis B vaccination has not been reported previously. This study found a prevalence of 54.7% in the POVP, which was lower but not statistically significantly different (p=0.08) from the pre-vaccination era (68.7% in the PRVP). It is possible that the prevalence of true OBI in South Africa was previously underestimated, thus
obscurring the actual impact of the universal hepatitis B vaccination programme on OBI. Determining the prevalence of seropositive OBI in the population is of high relevance to public health as it can be transmitted to susceptible individuals (Ocana et al., 2011; Raimondo et al 2013). In contrast, the potential for transmission from seronegative occult HBV infected individuals remains undetermined (Raimondo et al 2013).

Multiple studies on OBI have been conducted in South Africa mainly in HIV infected populations, where OBI was defined only by the detection of serum HBV DNA (regardless of the viral load) in the absence of HBsAg. In these studies, the prevalence of “OBI” (either true or “false”) has been shown to range from 22.1%-88.4% in the HIV infected population (Mphahlele et al, 2006; Lukhwareni et al, 2009; Firnhaber et al, 2009; Barth et al, 2011; Ayuk et al, 2013). In contrast, Bell et al (2012) examined true OBI in their population of HIV-infected patients from Southern Africa and found a prevalence of 6.7% (3/45) which is much lower than that found in the current study population (75.0% [9/12] in the POVP and 63.6% [14/22] in the PRVP). It is possible that the higher prevalence of OBI is a result of the health facility-based population studied, which is generally a high risk population compared to that in the Bell et al (2012) study which comprised participants presenting for initiation of antiretroviral therapy. In addition, the overall prevalence of HBV DNA in the HBsAg negative but anti-HBc positive (with or without anti-HBs), HIV-infected subset was also high; 91.7% (11/12) and 63.6% (14/22) in the POVP and PRVP respectively, compared to the 18.3% (33/180 [16/57 positive for anti-HBc only and
17/123 positive for both anti-HBc and anti-HBs] found by Bell et al (2012), possibly attributable to the difference in HBV DNA screening techniques used by either study. The HBV S gene was amplified for one isolate (PRE-9576) from a true OBI and no unique or signature amino acid variations were found. However, unique variations were identified in isolates from “false” OBI: S45P, P70H and V168A+P217L, although one isolate (PRE-8985) from a “false” OBI had the wild type HBsAg sequence. The cause of “false” OBI is often infection with HBV S gene variants with diagnostic escape mutations (Carman, 1997; Raimondo et al, 2008; Alavian et al, 2012).

The influence of the S45P and P70H variations on diagnostic escape cannot be deduced from this study. While a previous study by Kwange et al (2013) isolated the S45P variant from HBsAg positive donated blood, this variation is also known to be located within a cytotoxic T lymphocyte epitope at 28aa-51aa of the S protein. Variations within this region are associated with HBsAg undetectability (Tai et al, 1997; Alavian et al, 2012). The P70H variation found in isolate POST-4162 has been previously identified in vaccinated infants born to HBsAg positive mothers in China (Su et al, 2013). The V168A and P217L variations were observed within the same isolate (PRE-5603). This V168A variation has been reported previously in HBsAg carriers and has not been shown to alter the antigenicity of the protein (Weinberger et al, 2000; Ruiz-Tachiquín et al, 2007; Scheiblauer et al 2010). In contrast, the P217L substitution which has been reported in solely anti-HBc positive sera, is located within a T helper cell epitope at 215aa-223aa of the protein, and has been linked with diagnostic escape (Ducas et al, 1996; Weinberger et al, 2000; Kuzin et al
2012). It could be deduced then that the P217L substitution and not the V168A resulted in isolate PRE-5603 escaping HBsAg detection during serological testing.

Other variations of note included the M103I found in four isolates (POST-0225, POST-4028, PRE-1082 and POST-4288), all of which tested HBsAg positive. While M103I has been isolated in HBsAg positive infections it is typically isolated from HBsAg negative as well as solely anti-HBc positive sera (Weinberger et al, 2000; Panigraphi et al, 2010). The nonsense mutations Y200* and L206* located towards the end of the HBV S protein of isolates POST-5162 and PRE-1082 respectively, did not affect HBsAg detection as both isolates were from HBsAg positive HBV infections.

Variations within the MHR (110aa-165aa) of the HBsAg are reported to affect reactivity to anti-HBs (either host generated or those incorporated in diagnostic kits), allowing the virus to evade immune and vaccine response as well as serological detection (Carman, 1997; Echevarría and Avellón, 2006; Yokosuka and Arai, 2006; Alavian et al, 2012). While variations were observed within the MHR of study isolates (Table 2), the majority (K122R, N131T and T143S) were neutral substitutions. The A159P variation has been identified in HBV sequences from neighbouring Zimbabwe but its significance to the protein is unknown (Gulube et al, 2011). The E164D variation on the other hand has been linked with reduced binding of HBsAg to anti-HBs and is selected for by a Lamivudine resistance mutation (V173L) located within the overlapping polymerase open reading frame (Torresi, 2002; Torresi et al., 2002).
Finally, Kramvis and Paraskevis (2013) recently used a large dataset of full length HBV reference sequences – those included in this study – to show clear monophyletic groups for subgenotypes A1 and A2. We noted the same (data not shown). When the reference dataset was trimmed to the size (~681 bp) of the HBV S gene fragment amplified in this study, the ability to distinguish subgenotype A1 from A2 was not as obvious (Figure 1). Nonetheless, phylogenetic analysis of all 37 HBV S gene study sequences demonstrated the predominant HBV subgenotypes to be A1 (34 isolates) and A2 (2 isolates), in line with South African literature (Kimbi et al, 2004). The HBV genotype D is second to genotype A in terms of predominance in Southern Africa, with the dominant D subgenotype reported to be the D3 (Kimbi et al, 2004; Norder et al, 2004). However, recent studies have also isolated the subgenotype D4 from South African populations (references should appear here) as was the case in the current study, showing potential importation and resulting in further diversity in the HBV strains circulating within the country (Mayaphi et al, 2013). Subtype D4 has been reported in other African countries including Tunisia, Ghana, Somalia, Rwanda and Kenya (Borchani-Chabchoub et al, 2000; Norder et al, 2004; Candotti et al, 2006; Hübschen et al, 2009; Kwange et al, 2013).

In conclusion, this study reports the prevalence of true OBI, based on the definition from the Taormina expert meeting, in post-vaccination South Africa. Further investigations with similar study designs in more community-based populations are required in order to assess the actual impact of universal hepatitis B vaccination on true OBI in South Africa. In addition, HBV S gene variants including diagnostic escape mutants were identified particularly in the PRVP. Thus selection of vaccine
escape mutants following 17 years of a universal hepatitis B vaccination programme remains uncommon to South Africa.

**DECLARATION OF POTENTIAL CONFLICT OF INTEREST**

None

**ACKNOWLEDGEMENTS**

This work was supported by the National Health Laboratory Services (NHLS) Research Trust [grant number: GRANT004_94329] and the Poliomyelitis Research Foundation (PRF) [grant number: 11/74 (MSc)].

**REFERENCES**


Kuzin SN, Zabotina EE, Zabelin NN, Kudriavtseva EN, Samokhvalov EL, Borisova OV, Manina TA, Lisitsina EV, Kuzina LE, Malyshev NA, Lavrov VF. 2012. Heterogeneity of hepatitis B virus and diagnostic potential of modern test systems for the detection of HBsAg [Article in Russian]. Zn Mikrobiol Epidemiol Immunobiol 1:68-75.


