The Association between Hepatitis B Virus Longitudinal Genetic Variability and Clinical Outcome in Circumpolar Indigenous Populations

by

Kaarina Kowalec

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfillment of the requirements of the degree of

MASTER OF SCIENCE

Department of Medical Microbiology

University of Manitoba

Winnipeg

Abstract

The true prevalence of hepatitis B virus (HBV) infections in Northern Canada is underreported owing to the lack of occult-HBV studies. Clinical outcomes of HBV infections are variable; such that active disease is rare in Canadian and Greenlandic Inuit, yet hepatocellular carcinoma (HCC) is prevalent among genotype F (HBV/F) infected Alaskan Natives. The purpose of this study is to determine the true extent of occult infections in Northern Canada and the rate, nature and regional susceptibility of HBV genomic mutations among circumpolar indigenous populations.

The occurrence of occult infections in 700 archived serum samples from Northern Canada was determined by viral DNA amplification. In addition, HBV mutational analysis was performed on 15 indigenous peoples infected by one of 3 HBV genotypes (B6, D and F), which are associated with varying outcomes, including inactive liver disease and HCC. Phylogenetic analyses and genetic variation was investigated with full-length HBV genomes.

The results show 3.8% of 700 study subjects to be occult-HBV positive, with half of them infected with HBV/A. This study also reveals HBV/F strains contain deletions and substitutions associated with adverse outcomes, while HBV/D and B6 sequences lacked these mutations and contained mutational patterns associated with a benign outcome. The genetic diversity within the dominant and subclonal population levels are higher for HBV/B6 strains compared to HBV/D and HBV/F strains. The increased genetic variability found in virus associated with inactive disease may be indicative of an escape mechanism used by HBV to evade the host immune response or simply co-evolution between the virus and its host. These observations may demonstrate the

association between differing clinical outcomes and infecting genotype in circumpolar indigenous populations.

Acknowledgements

I would like to thank my Masters of Science supervisor, Dr. Carla Osiowy. Dr. Osiowy has provided to me the opportunities and knowledge required to progress my career in scientific research. She has played an indispensable role in my learning, in addition to performing exhaustive reviews of this thesis and for that I am forever grateful.

I would like to thank my two MSc supervisory committee members, Dr. Blake Ball and Dr. Gerald Minuk. They both provided wonderful guidance and assistance throughout this course of work to make this thesis the best it can be.

I would like to thank the Government of Canada and the Province of Manitoba for their financial support during this graduate work.

I would like to thank my lab members in the Bloodborne Pathogens and Viral Hepatitis section at the National Microbiology Laboratory, especially Elizabeth Giles and Jamie Borlang for their lab training, advice and continual support.

Last, but certainly not least, I would like to thank my husband, Max, for providing much encouragement, support and advice throughout this degree. This process would have been impossible without him. I would also like to thank my parents for providing me with guidance and support and to always try my best since day one.

Table of Contents

List of Tables

List of Figures and Illustrations

List of copyrighted material for which permission was obtained

Table 1. The clinical outcomes and geographical regions that have been associated with hepatitis B virus genotypes. Table reproduced with kind permission from Elsevier, Copyright (2010). Taken from [4].

Figure 2. Changes within the serological and clinical profiles of a patient acutely infected with hepatitis B virus. Figure reproduced with kind permission from Elsevier, Copyright (2009). Taken from [5].

Figure 3. Structural organization of the hepatitis B virus genome. Figure reproduced with kind permission from Nature Publishing Group, Copyright (2005). Taken from [3].

Figure 4. The prevalence of hepatitis B surface antigen and global geographic distribution of hepatitis B virus genotypes. Figure reproduced with kind permission from John Wiley and Sons, Copyright (2010). Taken from [6].

List of Abbreviations

AA	Amino Acid
ALT	Alanine Aminotransferase
Anti-HBc	Antibodies to Hepatitis B Core Antigen
Anti-HBs	Antibodies to Hepatitis B Surface Antigen
Anti-HBe	Antibodies to Hepatitis B e Antigen
Anti-HBx	Antibodies to Hepatitis B X Antigen
BCP	Basal Core Promoter
cccDNA	Covalently Closed Circular DNA
dN	Nonsynonymous Change
DNA	Deoxyribonucleic Acid
dNTPs	Deoxyribonucleotides
DR1	Direct repeat 1
dS	Synonymous Change
EHN1	Enhancer 1
GTR	General Time Reversible model
HBV	Hepatitis B Virus
HBcAg	Hepatitis B Core Antigen
HBeAg	Hepatitis B e Antigen
HBe-SC	HBeAg seroconverted
HBIg	Hepatitis B Immunoglobulin
HBsAg	Hepatitis B Surface Antigen
HBx	Hepatitis B X Protein
HCC	Hepatocellular Carcinoma
HCV	Hepatitis C virus
HDV	Hepatitis D virus
HIV	Human Immunodeficiency Virus

HLA	Human Leukocyte Antigen
IFN	Interferon
IDU	Injection Drug User
IU	International Units
LB	Lysogeny Broth
NCBI	National Centre for Biotechnology Information
NML	National Microbiology Laboratory
NNI	Nearest Neighbour Interchanges
nt	Nucleotide
OBI	Occult hepatitis B virus infection
ORF	Open Reading Frame
pgRNA	Pregenomic RNA
POL	Hepatitis B virus polymerase protein
PreC	Pre Core Region
PreS	Pre Surface Region
PCR	Polymerase Chain Reaction
RC	Relaxed circular
RNA	Ribonucleic Acid
RPM	Rotations Per Minute
RT	Reverse transcriptase
RT-PCR	Real Time PCR
SPR	Subtree Pruning and Regrafting Algorithm
SOC	Super Optimal Broth with Catabolite Repression
TA	A1762T/G1764A mutation
TAE	Tris-Acetate-EDTA buffer
tRNA	Transfer RNA
WHO	World Health Organization

Introduction

Hepatitis B virus (HBV) infections result in over 600 000 deaths each year and 350 million chronic infections worldwide [7]. Infections with HBV are considered endemic within the indigenous populations of the western circumpolar region (Fig. 1). This region includes Alaska, northern Canada and Greenland [8]. The prevalence of active HBV infection (based on HBV surface antigen [HBsAg] positivity) within Canadian Inuit is 20 times that of Caucasians inhabiting southern Canada [9]. Several reports also estimate the risk of HBV exposure is 5 times higher within these Inuit residents, when compared to southern Canadians. A recent study demonstrated Greenlandic Inuit have a higher seroprevalence of active HBV infection (20%) than other northern European countries (<0.5%) [10]. In addition, amongst all ethnic groups in the US, Alaska Natives had the highest prevalence (6.4%) and incidence of HBV [9]. Implementation of routine universal HBV vaccination during the mid- to late-1980s is expected to have lowered the incidence rate in Alaska and Canada [11-12].

Infections occurring in the first years of life pose the greatest risk of developing into a chronic HBV infection [5]. The risk of progression to chronic HBV infection is 90% in infants infected at birth, while infection up to 5 years of age is associated with a 30% risk [5]. Prior to routine vaccination strategies [11-12], vertical transmission was the primary route of HBV transmission among Canadian Inuit and Alaska Natives [9], resulting in endemic rates of infection. In contrast, sexual contact is a more likely route of HBV transmission in Greenland, owing to the reported increase of infection rates in young adults [9]. Furthermore, there is no routine childhood HBV vaccination program available in Greenland [13] to prevent the cycle of new infections. Despite HBV vaccination

programs effectively reducing the rate of new infections among Alaska Natives, Canadian Inuit and First Nations populations of the Arctic [11], the prevalence rate is expected to remain high for generations to come. Such high prevalence rates may result in a significant burden of disease associated with HBV infection in these regions.

There exists an association between distinct clinical outcomes and infection with a specific HBV genotype within each western circumpolar indigenous population. For example, a significant association exists between genotype F (HBV/F) and the development of hepatocellular carcinoma (HCC) [14] and between HBV genotype D (HBV/D) and HBV-associated vasculitis amongst Alaska Natives [15]. In contrast, no clinical or biochemical evidence of active liver disease is found in Canadian or Greenlandic Inuit infected with genotype B6 (HBV/B6) [10, 16] or in Greenlandic Inuit infected with HBV genotype D (HBV/D) [10]. The clinical outcome of HBV-infected Canadian Arctic First Nation peoples, in which genotype D predominates [17], is not well understood.

This curious association between a specific HBV genotype and a clinical outcome in this ethnic population raises the question of the role of the virus genotype in HBV infection and disease. Our understanding of HBV clearance and disease manifestation during infection is related to the host immune response [5]; however, accumulating evidence now suggests a relationship between HBV genomic characteristics and clinical outcome [18-20]. Thus, the present study aims to explore the genetic characteristics of different HBV genotypes to elucidate the association of infecting genotype with clinical outcomes in circumpolar indigenous populations. Part of this exploration also includes understanding the prevalence of inapparent, or occult, HBV infection within a Canadian

Arctic First Nations population. Molecular characterization of virus resulting in occult infection provides further insight into the role of HBV in clinical outcomes experienced by indigenous populations of the western circumpolar Arctic.

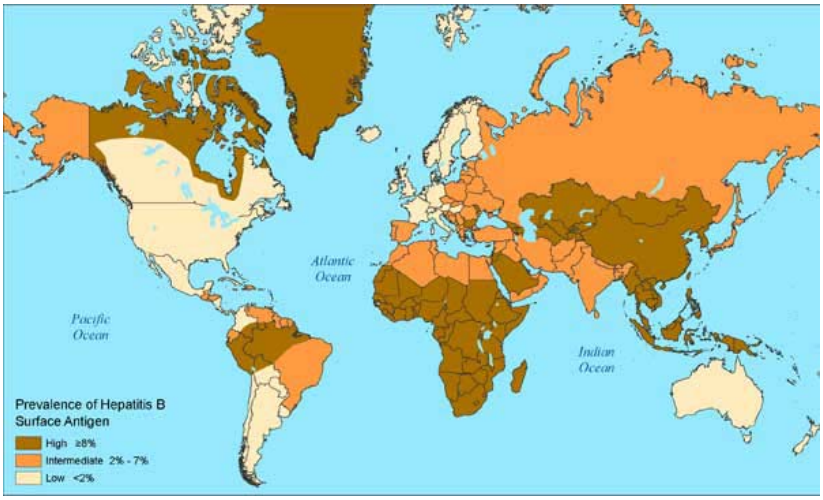
1. Transmission routes and serological markers of infection

Transmission of HBV occurs either through percutaneous or mucosal routes such as injection drug use, sexual contact or vertical transmission from mother-to-child [11]. Transmission via sexual contact is frequent in men who have sex with men and individuals with multiple sexual partners [21]. Additionally, infected health-care workers present a potential source of HBV infection to their patients via percutaneous transmission, although vaccination of health-care workers reduces this risk [22]. The risk of health care worker-acquired transmission remains high in developing countries (Fig. 1), owing to the lack of immunization programs, sufficient sterilization, and infection control practices [11].

Mother-to-child transmission can occur either in-utero or during the birthing process. The child's risk of infection is 5-20%, if the mother is HBsAg-positive [21]. The HBsAg is the most abundant of all proteins produced by HBV and indicates active infection [5]. Vertical transmission is highly associated with developing chronic infection and is thought to be the reason for the Northern indigenous population of Canada having one of the highest HBV prevalence rates in the country (based on HBsAg positivity) at 3-5% [9]. Approximately five to ten percent of Canadians show a marker of previous HBV exposure (antibody to HBV core antigen [HBc] – anti-HBc) [23]. Anti-HBc IgG is observed during chronic infection but also persists in exposed individuals who have resolved

infection [5]. Anti-HBc IgM antibody normally indicates a recent, acute HBV infection, but can be observed in chronically infected patients during an acute flare [24]. Antibody to HBsAg (anti-HBs) indicates a previous, resolved infection [5] or is often observed following clearance of HBsAg during the inactive phase of chronic infection (anti-HBs seroconversion) [4]. Additional markers, such as HBV "e" antigen (HBeAg) are required for definitive diagnosis of acute and chronic infections. When a pregnant mother is HBeAg-positive, her child's risk of infection is as high as 90% [21]. The presence of HBeAg past 10 weeks and/or HBsAg past 6 months indicates a likely progression to chronic infection [5].

Figure 1. Worldwide prevalence of hepatitis B virus surface antigen (HBsAg) (2005). This figure estimates the low, intermediate and high areas of HBsAg prevalence in 2005. Source: Centre for Disease Control, Health Information for International Travel 2010. Atlanta: U.S. Department of Health and Human Services, Public Health Service, 2009. Note: For Northern Canada, the prevalence of HBsAg is approximately 4%, not ≥8% as stated.



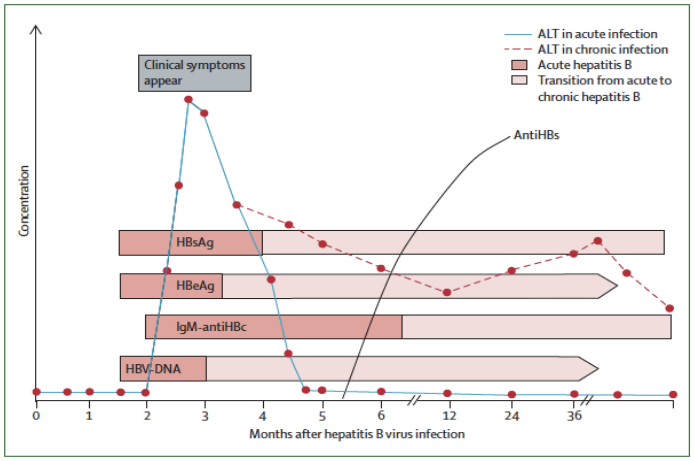
2. Natural history of hepatitis B virus infections

2.1 Acute hepatitis B infections

As previously mentioned, clinical outcome may be associated with the infecting genotype. Acute infection with HBV genotype D2 (HBV/D2) has been associated with self-limiting hepatitis [25]. Acute infections with HBV initiate after an incubation of 45-180 days, with the first and most abundant serological marker (HBsAg) appearing 1-6 weeks before clinical symptoms (Figure 2) [5]. HBeAg and HBV DNA are both markers of viral replication, and appear shortly after the HBsAg. The onset of clinical hepatitis coincides with elevations in host alanine aminotransferase (ALT) levels, as indicated by liver biochemical testing. HBV-specific CD8 T cells recognize infected hepatocytes to control the virus, but also recruit non-specific inflammatory cells causing hepatic injury. The innate and adaptive immune responses initiate removal of HBV DNA, followed by cytolysis of hepatocytes and necrosis [5]. Researchers believe a multivalent CD4 and CD8 response clears an acute HBV infection and is accompanied by the production of antibodies to HBsAg (anti-HBs) [5]. As clearance of acute HBV infection is mediated via the immune system, antiviral treatments are not normally employed. Those with ineffective CD4 and CD8 responses may progress to persistent, chronic HBV infections [5].

Figure reproduced with kind permission from Elsevier, Copyright (2009). Taken from [3].

Figure 2. Changes within the serological and clinical profiles of a patient acutely infected with hepatitis B virus (HBV). The figure also displays the alanine aminotransferase (ALT) levels of acute and chronic HBV infections. The dark pink bars demonstrate the levels of the associated HBV marker during acute infections, whereas the light pink bars show these markers during chronic infections.



2.2 Chronic HBV infections and treatment

Chronic HBV infections follow a complex natural course of four phases; immune-tolerance, immune-active, inactive and HBsAg clearance [4]. For chronically infected patients, interferon and reverse transcriptase (RT) inhibitors are the two treatment categories currently available [5]. The interferons (IFN) act by stimulating T-cells during the immune response [7]. The general antiviral effects of IFN afford a low risk of developing resistance, but are generally, less efficient than the nucleos(t)ide inhibitors at reducing HBV DNA levels [26]. The second class of HBV antivirals, RT-inhibitors, inhibit the HBV DNA replication enzyme (reverse transcriptase, RT) by incorporating nucleotide analogs [7]. A patient chronically infected with HBV may or may not experience

the first immunotolerance phase, as it primarily occurs in persons infected perinatally from HBeAg-positive mothers [4]. Normal levels of ALT, very high HBV DNA and minimal liver inflammation characterize this phase. The lack of immune recognition towards HBV allows this phase to persist for potentially more than 40 years. Those infected with HBV later in life normally do not experience the prolonged immunotolerance phase [4].

The immune active phase is characterized by HBeAg-positivity, slightly lower HBV DNA levels, elevated ALT and active liver disease [4]. Although weak, a cytotoxic T cell response activates and eventually HBeAg seroconversion (anti-HBe-positivity) occurs, as this phase progresses [4]. Following seroconversion to anti-HBe positivity (HBe-SC), chronic HBV infection may proceed in 3 potential ways. A reversion to HBeAg seropositivity occurs in 10-40% of patients and is more commonly observed in genotype F-infections than HBV/D or HBV/B infections [27]. This HBeAg seroreversion may result in a subclinical flare of hepatitis. In the second possible outcome, 20% of patients remain immune-active with slightly higher levels of HBV DNA [4].

The third and most likely phase for patients is the "inactive HBV phase" [4]. Normalized levels of ALT, lower HBV DNA and HBeAg negativity characterize this third phase, where most will remain for life. Usually, a strong HBV-specific cytotoxic T-cell response is required for patients to remain here, and this potentially benign carrier status is also associated with infections with HBV genotype B6 (HBV/B6) [4]. Injury incurred to the liver from previous stages of chronic infection may improve and even reverse. The patients unable to mount a successful cytotoxic T-cell response may reactivate to the immune active phase [4]. Diagnosed using ultrasonography or liver

biopsies, liver cirrhosis and/or HCC may occur in patients displaying cycles of inactivation followed by reactivation. Patients infected with HBV genotype C, D2 and D3 all have increased rates of cirrhosis and/or HCC, when compared to other genotypes [28] and a risk of such outcomes requires that many patients be followed closely, with monitoring of ALT levels for life [4].

Clearance of HBsAg differentiates the fourth and final phase where most, but not all patients develop anti-HBs [4]. The estimated rate of annual HBsAg clearance in chronic HBV infections is between 0.5-0.8% per year, with older age associated with clearance of HBsAg. The risk of cirrhosis in chronically infected patients with HBV is approximately 10-30% [29]. The risk of both cirrhosis and HCC decrease upon spontaneous HBsAg clearance, however; there is evidence the risk of HCC development remains significant, especially in patients infected with HBV genotype F (HBV/F) [4, 14]. The risk of HCC following HBsAg clearance is higher for males >40 years, females >50 years and with a positive family history for HCC [4].

2.3 Mechanisms of HBV-associated liver disease

Stages of inflammation and fibrosis of the liver, and increased liver cell DNA synthesis characterize the development of cirrhosis. Patients with active cirrhosis may progress to HCC, with approximately 1-3% of cirrhotic cases developing liver cancer per year, demonstrating that HCC generally, yet not exclusively, develops in cirrhotic patients [30].

While co-infections with the hepatitis D virus (HDV) [31] or the human immunodeficiency virus (HIV) [32] play a role in progression from chronic HBV infection

to HCC, there exist three prevalent hypotheses in the mechanism of HBV-related HCC [29]. Firstly, research shows HBx enhances viral replication and possesses direct transforming activities [33]. A further potential mechanism of HBV-related liver cancer is the immune-mediated chronic injury path [34]. Years of cell destruction and regeneration cycles from intermediate immune responses leads to steatosis, cirrhosis and fibrosis and eventually HCC. More likely, this second mechanism occurs in the absence of HBV host integration and HBx expression [34]. Thirdly, integration of HBV DNA is found in the chromosome of liver tumor cells, causing deletions leading to chromosomal instability and thus, cancer [29]. Insertion of HBV DNA within proliferative- and differentiation-associated genes also contributes to HCC development in a small subset of cells. The actual incidence of HBV insertion is not known, although it is acknowledged HBV varies its own genomic makeup quite frequently [5]. The extent of genetic variation is believed to largely depend upon the HBV replication cycle [5].

3. Genotypes and genomic variability

3.1 Genomic organization

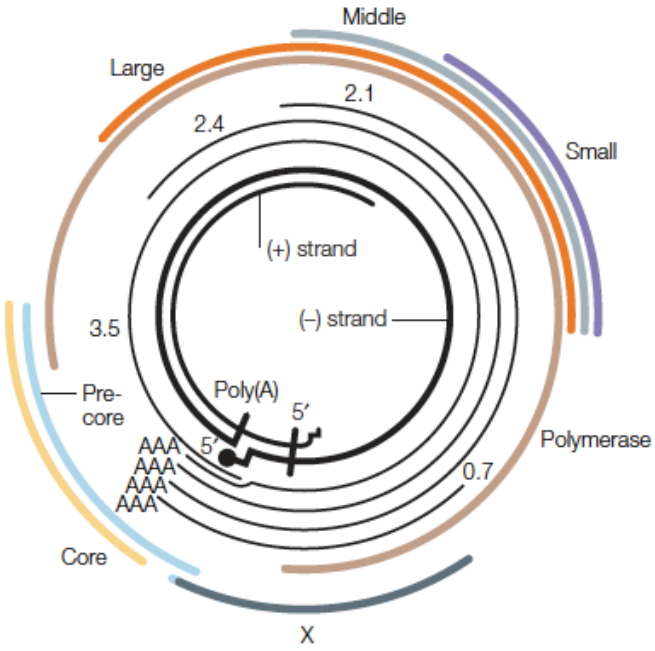
The infectious form of the 3.2 kb partially double-stranded DNA genome of HBV is given the term relaxed circular DNA (RC-DNA) [7]. The virus is preferentially tropic for hepatocytes and upon entry and transport of the viral genome to the nucleus, a 5' RNA oligonucleotide is removed from the partial positive strand [7] and the viral polymerase is removed from the 5'-end of the minus strand [7]. Thereafter repair and ligation of both strands allows formation of cccDNA (covalently closed circular DNA). The host cellular RNA polymerase II then transcribes multiple genomic and subgenomic RNA species

from the cccDNA, including the pregenomic RNA (pgRNA). The pgRNA contains greater than full-length HBV sequence and serves as the mRNA for the DNA polymerase having reverse transcriptase activity, HBV core antigen (HBcAg) and progeny genomes [35].

The four overlapping open reading frames (ORF) of this enveloped virus require several replication signals and enhancer sequences to transcribe the various proteins from the small number of ORFs (Figure 3). The "S" ORF codes for three proteins (small, medium, and large surface proteins) [7], which along with host lipid, compose the viral envelope. The least common surface protein, large surface antigen, is expressed from the transcript preS1, whereas the medium and small surface proteins, transcribed from preS2 and HBsAg mRNAs, respectively, make up the majority of the envelope. HBsAg is also secreted in excess as self-assembled subviral particles composed of protein and host lipid [36]. This forms the majority of antigen detected in HBsAg serological detection assays and is the basis of the recombinant HBV vaccine [36]. The core ORF has two in-frame start codons, both using the same stop codon [7]. The first codon (preC) produces a post-translationally modified, secreted protein (HBV "e" antigen, HBeAg) and the second start codes for the structural capsid protein (HBV core protein, HBcAg). The function of HBeAg is not well understood, as it is not required for replication, assembly or infection; yet is a marker of active infection. HBeAg also plays a role as an immunoregulatory protein that downregulates the immune response to HBcAg [37]. The last two reading frames code for the X protein (HBx), which may play a role in HCC progression [38] and the polymerase (POL) ORF coding for the DNA replication enzymes, RT [7].

Figure reproduced with kind permission from Nature Publishing Group, Copyright (2005). Taken from [3]

Figure 3. Structural organization of the hepatitis B virus (HBV) genome. This is a graphic representation of the virus, with the designated open reading frames, as the core, pre C (pre-core), X, PreS1, PreS2, S (surface) and polymerase. Located in the inner portion is the partially double stranded DNA genome of HBV.



3.2 Genomic classification

Kramvis *et al* define a genotype as: "the genetic constitution of an organism" [39].

Sequence classification of the HBV virus has led to the differentiation of 8 different genotypes, named A through H. This strain separation is likely due to the relatively rapid evolutionary rate of HBV (approximately 10^{-5} nucleotide substitutions per site per year), in comparison to other DNA viruses [16], and the formation of specific phylogeographic lineages [40]. A new genotype of HBV must adhere to the following guidelines: 8% sequence divergence over the complete genome of approximately 3200 nucleotides, and

>4% within the HBsAg (approximately 681 nucleotides) [7]. Subgenotypes are further classified by 4-8% sequence divergence within a group.

3.3 Divergence of genotypes

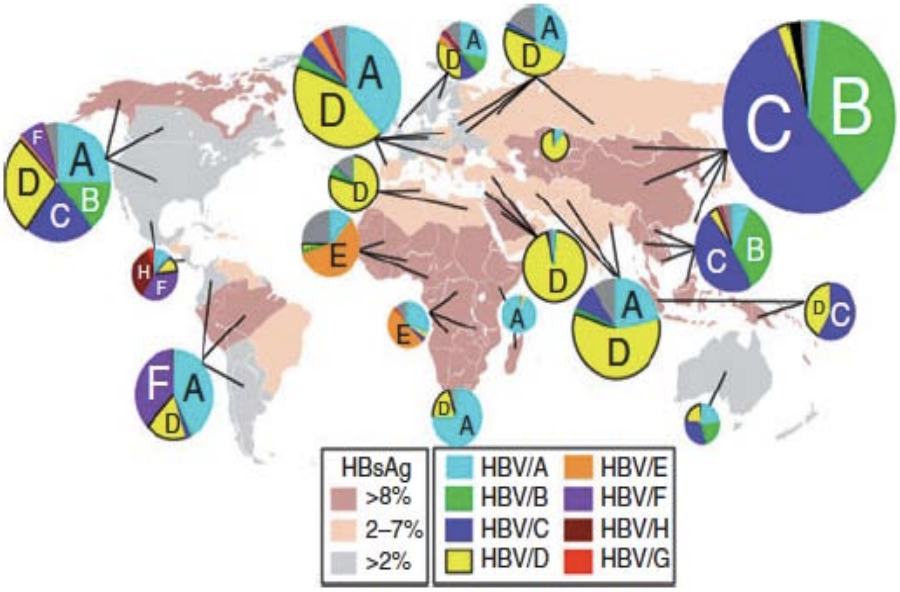
The characteristic geographical distribution of the eight HBV genotypes, as shown in Figure 4, aid in molecular epidemiological studies and further characterization of HBV.

A small insertion located in the HBcAg carboxyl terminus region distinguishes genotype A (HBV/A) from the other seven genotypes [6]. This genotype has five subgenotypes: A1-A5 and can be found in the America's, Asia, Africa and parts of Europe. Genotype B has seven subgenotypes (B1-7) and is prevalent in Japan, Asia and the Arctic. With the exception of HBV/B1 and B6, all subgenotypes of B contain a genotype C recombination within the core region of the genome [6]. Genotype C is distributed throughout Asia and Australia, with its five subgenotypes [6]. Genotype D is found worldwide and contains a characteristic deletion within the preS1 region. Genotype E is most closely related to genotypes A and D and is commonly found in coastal African and European countries [41]. Genotype F (HBV/F) has four subgenotypes (F1-F4) and has further subdivisions within F1 and F2 [6]. HBV/F is commonly found in the Aboriginal populations of Alaska and South America, as well as some newly identified isolates from Japan. Genotype G contains stop codon mutations within the pre-core (preC) coding sequence and infection almost always occurs as a co-infection with other HBV genotypes. This is evident as genotype G is unable to produce HBeAg, yet patients infected with this genotype are often HBeAg-positive. Genotype G is found within Europe, the United States and most recently, Canada [42]. Finally, genotype H is

limited to North and Central American populations and is most closely related to HBV/F [6].

Figure reproduced with kind permission from John Wiley and Sons, Copyright (2010). Taken from [6].

Figure 4. The prevalence of hepatitis B surface antigen (HBsAg) and global geographic distribution of hepatitis B virus (HBV) genotypes. This figure was produced from a literature review of 300 articles. The prevalence of HBsAg is also shown, based on 2005 estimations from the Centers for Disease Control. A pie chart was produced for each area where research has been conducted and the corresponding prevalence of each genotype is shown. Note: For Northern Canada, the prevalence of HBsAg is approximately 4%, not ≥8% as stated.



An infection with both genotype G and A is an example of a co-infection with multiple HBV genotypes [28]. When more than one genotype is prevalent in a region, a co-infection is more likely and can inevitably lead to new recombinants. It is believed a

recombination between ancestral genotype B and genotype C strains gave rise to subgenotypes B2-B5 and B7 [28]. Other examples of common recombination events include; genotype A and D, or genotypes C and D. The HBV genomic regions of nt1-200 and nt1600-approximately 3200 may be "hot-spots" for recombination [28]. How exactly recombination occurs and the clinical significance associated with infection with a recombinant virus remains unknown [28].

4. Genetic variation of HBV and the association with clinical outcome

4.1 Mechanisms behind hepatitis B viruses' genetic variation

Research supports the hypothesis of multiple factors contributing to the inherent genetic variation of HBV. The overlapping ORFs (Figure 3) limit lethal mutations, while non-lethal mutations are quite prevalent [39]. The use of RT during replication of the virus is a major reason for introducing these mutations. This DNA replication enzyme lacks proofreading capabilities, and introduces mutations at a rate comparable to retroviruses, such as HIV [39]. In addition, the host's immune response, and external or environmental factors, such as antiviral treatment and immunization, all place selective pressure on HBV to continually mutate and avoid removal either via the host or via therapeutics. These mechanisms all contribute in many ways to the inherent genetic variation of HBV, and to the reason for multiple genotypes.

4.2 Evidence for the association between genotypes B6, D and F and clinical outcome

Research involving HBV places a high importance on viral genotyping, because the genotype may be capable of predicting the clinical outcome. Many of these studies are restricted solely to the areas of Asia and Europe endemic for HBV [43]. There are many studies describing an association with clinical outcome, although some remain inconclusive, as in the case of HBV/A infection and HCC development [28]. Yet other research shows significant associations, as in the example of increased rates of HCC and cirrhosis found with two subtypes of HBV/C [28].

Subgenotype B6 (HBV/B6) is the only subgenotype of B with no association with adverse outcomes [28]. Studies have shown earlier HBeAg seroconversion (HBe-SC) is more common in genotype B infections in comparison to genotype C infections [44]. There is an association between HCC and chronic infection with HBV/B2 in individuals younger than 50 years lacking cirrhosis [45]. In comparison to HBV/C2, a common double mutant is found less often in HBV/B2 [44] and many researchers hypothesize an increased risk for adverse outcomes involves multiple mutation sites [46]. HBV/B1 is associated with a higher risk of fulminant hepatitis and HCC in persons aged greater than 34 years [47-48]. There are higher rates of HCC in those infected with HBV/B3-B5 and B7, as shown in Table 1 [4]. The common TA (A1762T/G1764A) double mutation in HBV/B isolates is related to the development of acute-on-chronic liver failure [49].

HBV genotype D (HBV/D) infection appears to progress through a complex natural history dependent on the subgenotype. In comparison to HBV/A-infected patients, HBV/D infections have lower HBsAg clearance rates [50]. Vasculitis associated

with acute HBV infection is more common in patients infected with HBV/D, than genotypes A and F [15]. Acute self-limiting hepatitis is common in HBV/D2 infections [25]. In addition, most D subgenotypes, with the exception of D4 have been associated with progression to HCC (Table 1) [4].

Studies into HBV genotype F1 infection show associations with developing HCC [4]. In comparison to genotypes A2, B6 and D, a significantly higher proportion of Alaska Natives infected with HBV/F1 developed HCC [14]. Interestingly, this association in HCC and HBV/F1-infected patients depends upon the patient being much younger than HBV/A2, C2 and D-infected patients. An association exists between HDV and HBV genotype F2 or F3 co-infections and fulminant hepatitis [4]. These studies all provide evidence for an association between the infecting genotype and clinical outcome. In addition to the aforementioned eight genotypes, genetic variation also results in mutations, many of which are well characterized with respect to an association with clinical outcome in HBV infections [51].

Genotype	Geographic Region	HBV Disease Association
A1	Central and eastern sub-Sahara Africa	HCC in young males
A2	Northern Europe	HCC and cirrhosis in older persons
A3	West Africa	HCC
B1	Japan	HCC and cirrhosis in older persons
B2-5, B7	East Asia	HCC and cirrhosis occur at younger age than B1
B6	Indigenous Populations in Alaska, Canada, Greenland	No serious sequelae identified to date
C1	Vietnam, Thailand, Myanmar, Indonesia	High rates of HCC and cirrhosis
C2	China, Taiwan, Korea, Japan	High rates of HCC and cirrhosis
C3	Pacific Islands (Micronesia, Melanesia, and Polynesia)	High rates of HCC and cirrhosis
D1	Europe, Middle East, Egypt, India, Asia	HBeAg-negative chronic hepatitis/ cirrhosis and HCC
D2	Europe, Japan	HBeAg-negative chronic hepatitis/ cirrhosis and HCC
D3	Europe, Asia, South Africa, United States	HBeAg-negative chronic hepatitis/ cirrhosis and HCC
D4	Australia, Japan, Papua New Guinea	Not studied
E	West Africa	Not studied but area of high incidence of HCC
F1	Alaska, Argentina, Bolivia	HCC in young patients in Alaska
F2	Venezuela, Brazil	Fulminant hepatitis with HDV co-infection
F3	Venezuela, Columbia, Panama	Fulminant hepatitis with HDV co-infection
F4	Argentina	Not studied
G	France, United States, Vietnam	Usually found in co-infection with genotype A; increased association with acute hepatitis, liver fibrosis, and HCC in Vietnam
H	Mexico, Nicaragua, California	Not studied

Table 1. The clinical outcomes and geographical regions associated with hepatitis B virus (HBV) genotypes. This table was produced from a literature review of 79 articles.

4.3 Specific mutations associated with clinical outcomes

The commonly observed mutation at nt 1896 (A1896) results in a premature stop codon within the preC region leading to the loss of expression of HBeAg [51]. A recent study has found that this mutation provides a selective advantage to the virus as a CTL escape mutant following breakdown of immunotolerance [37]. Escaping CTL targeting of host

hepatocytes provides another possible mechanism of viral persistence. The mutation at nt1896 is the most common preC variant and is found most often in genotype D and rarely in genotype F strains [18]. Although initially believed to be associated with only fulminant outcomes, this mutation is also found in benign cases [18]. Recently, a study comparing viral markers in chronic infection showed the A1896 mutant was more common in benign carriers than in cirrhotic and immune-tolerant cases [52]. The G1899A mutation was also shown to be significantly associated with cirrhotic liver cases [19].

The core region of HBV possesses high genetic variability as it is a major antigenic site [7]. There are many mutations within this region associated with HCC, however; two are observed fairly frequently, T2170C and T2441C, especially in genotype C infections [20]. There have also been significant increases of liver fibrosis in patients with core protein mutations [53].

The double mutant, A1762T/G1764A (TA), is the most common variation within the basal core promoter (BCP) region. The presence of this mutation in association with certain genotypes has clinical relevance, as it is associated with the development of HCC [4]. The mutation on its own, regardless of genotype, is associated with lower production of HBeAg and preC mRNA and is highly reported in chronic HBV cases. There are many other mutations within the BCP/X region that are associated as independent or dependent risk factors for HCC such as G1613A, G1653T, A/T1752C, T1753V, C1766T and T1768A [48, 54-55].

Within the PreS region, there are various mutations such as T31C, T53C, C2964A, C3116T, and C7A that are independently associated with HCC development [20]

and cirrhotic cases [46]. Mutations within the HBsAg region cause a variety of problems for diagnostic and immunization outcomes. The virus may escape some diagnostic tests as many capture antibodies, including those produced in response to the HBV vaccine and may not recognize a PreS/HBsAg mutant [56]. Thus it has been speculated that acquiring a mutation within the HBsAg may cause protection conferred through vaccination to be lost [18]. In vaccine- and immune-escape HBV isolates, the most common variant is a substitution of glycine to arginine at amino acid 145 (G145R) within the antigenic determinant region [18].

The Enhancer I (ENHI) region, located upstream of the HBx protein, regulates expression of this protein [57]. A study by Cho *et al* (2010) shows two sites within the ENHI, G1053A and G1229A, are independently associated with HCC development in cirrhotic patients infected with HBV/C2 [57]. These substitutions may cause transactivation of the ENHI and contribute to increased levels of HBx, the protein implicated in HCC development.

Although a rare occurrence, mutations within the RT protein are observed during long-term antiviral treatment [58]. Mutations within this protein may cause resistance towards antivirals that target the RT protein and will either lower or completely prevent an antiviral treatment from effectively eliminating an HBV infection. When treating with lamivudine, a nucleotide analog, HBV may attain the rtM204V/I mutation, and render the drug inactive [59]. Developing resistance towards an antiviral treatment is one of most common reasons for treatment failure and presents a problem for many physicians and patients [58]. Additionally, HBV may possess these resistant mutants in smaller subsets of the genomic population, at the clonal level. Upon treatment, clonal strains with antiviral

resistance emerge due to superior fitness levels over those lacking these substitutions [60]. Thus, if these resistance mutants do not exist as the most prevalent entity within a patient, it can be difficult to detect them with diagnostic assays. These subclonal populations are collectively referred to as the quasispecies population.

4.4 Quasispecies populations of HBV

The use of an error-prone DNA polymerase by HBV contributes to the genetic variations leading to its existence as a quasispecies (QS) population [61]. A QS population implies the total intrapatient virus population exists as a spectrum of variants, which possess varying levels of fitness, depending on the environment [61]. This is also observed with other viruses, such as HCV and HIV [62]. An example of this environmental effect is observed with drug therapies that contribute to an extensive clonal population and provide a relapse reservoir following antiviral treatment [63]. Studies have found the intrapatient population diversity to either enhance or diminish the replicative ability of the virus, depending on mutations within specific regions of the viral genome [38]. The clonal population of the RT region is susceptible to antiviral treatments, making it a good candidate for investigating QS. Recently, an investigation into the contribution of QS diversity to drug therapy responses showed the QS population to differ between drug-responders and non-responders, with the latter being more diverse [61]. A more diverse subclonal population is an advantage to HBV, as it could allow evasion of antiviral treatments.

The HBsAg reading frame completely overlaps with that of the POL gene, as shown in Fig. 4. Thus, a non-responding patient undergoing nucleoside analog treatment

and possessing a more diverse POL QS population may also possess a more diverse HBsAg QS, when compared to a treatment responder [61, 64-65]. Changes within the HBsAg may lead to a reduction in recognition by the wild type-derived antibody of this envelope protein [64]. Massively parallel ultra deep pyrosequencing, a method successful in detecting minor, yet clinically relevant variants in HIV [66], detected a premature stop codon within a minority of the HBsAg QS in nucleotide analog-treated HBV-infected patients [65]. This premature stop codon may result in truncated HBsAg and/or RT proteins; however these variants more than likely represent a small subset of the QS population, owing to the requirement for a functional RT protein. However, dominance of other HBsAg mutant populations allowing for functional replication, such as the G145R mutant, exist and are associated with immune escape [67]. These mutant populations have been implicated in the loss of detection of HBsAg by serological assays, thus potentially resulting in a diagnosis of "occult" HBV infection.

4.5 Occult-HBV infections

Occult HBV infections (OBI) are characterized by the lack of detectable HBsAg in the presence of HBV DNA [68]. Often the only viral marker detected in OBI is antibody to the core protein ("anti-HBc only") [69]. Regions of the world at a higher risk for wild-type (HBsAg-positive) infections also have higher rates of OBI [69]. Blood transfusions [70] and orthotopic liver transplants (OLT) [71] are examples of how an inapparent HBV infection may be transmitted, with 17-94% of anti-HBc-only OLT donors transmitting these atypical HBV infections [71]. Few studies reporting the prevalence of OBI in Canada exist, although there are some estimates within northern Canadian Aboriginal communities [72].

A community-based study from 2005 found OBI in 10% of a northern Canadian Inuit population [72].

Why occult-HBV infections are HBsAg negative, despite the persistence of HBV DNA in the blood remains unknown [70]. HBsAg mutants may not be detected by monoclonal anti-HBs antibodies within diagnostic assays and virus carrying mutations preventing the expression of HBsAg will fail to be detected by all HBsAg serological assays. However, research shows the majority of OBI cases do not result from HBV genetic variation [70]. Three theories exist for the occurrence of occult-HBV. First, host suppression of both viral replication and HBsAg expression may cause these asymptomatic infections. Another theory for the existence of OBI involves co-infection with HCV, owing to HCVs strong ability to inhibit HBV replication [70]. The final theory states host epigenetic mechanisms possess a role in regulating HBV replication and transcription through deacetylation and methylation [70].

The clinical role of OBI lies within disease progression and the risk of reactivation [68]. The risk of occult HBV-infected patients having a reactivation of HBV infection and progressing to a severe or fulminant course of hepatitis does exist [73]. Treatment with chemotherapies inducing an immunocompromised state poses a high risk of reactivation upon reconstitution of the immune system [73]. Following recovery, liver damage appears in these patients via T-cell mediated inflammation. Co-infections with HCV and OBI pose a risk of accelerated progression to cirrhosis [68]. Although more studies are needed to define the role of OBI in HCC patients, researchers have determined the prevalence of OBI among HCC patients to be between 22-87% [74]. The

development of HCC in OBI is believed to act by similar mechanisms as the traditional form of HBV [68].

Rationale

The true extent of HBV infections in Northern Canada is unknown and may be under-reported, as current data is based upon HBsAg positivity. The potential of occult infections should not be ignored, as it may lead to a higher prevalence and increased burden of HBV-associated disease within Northern Canada. Furthermore, genetically characterizing HBV isolates could allow for the discovery of potential markers for adverse outcomes in infected patients. This area has been studied to some extent, but there are still many points to consider.

For example, investigations comparing multiple genotypes and their association to the development of cirrhosis and/or HCC are difficult because of the specific geographic, and thus ethnic, distribution of the 8 HBV genotypes. Studies from Taiwan, Japan and other Asian countries, where only two genotypes, B and C, are found within the population are an example of this [75-77]. Discerning the contribution of genotype to clinical outcome is even more difficult since both genotype B and C are associated with benign and adverse outcomes. As well, studying HBV from patients who inhabit large, multi-ethnic cities, such as Tokyo and Beijing, may not control for false-positive associations between genotype and clinical outcome due to the potentially large overlap between host geographic factors and genotype [28]. Finally, some researchers recommend longitudinal studies be done to study the true relation between viral variation and clinical outcomes [70]. However, most studies still focus on single time point samples or at best, 3-

year paired samples [78] likely owing to the difficulties in attaining longer-term paired serum samples.

These points all raise questions to be answered: What is the true-extent of HBV infections within Northern Canada, specifically the prevalence of occult HBV? What role does viral genetic variation and mutational rate play in the observed association between clinical outcome and HBV genotype?

This study aims to provide evidence in answering these puzzling questions. To do so, seven hundred archived serum samples from a Northern Canadian population are obtained. Here we will study only HBsAg-negative subjects from a region previously determined to have a high HBsAg-positivity rate [79] as these subjects represent a potentially large reservoir of occult HBV infections. Additionally, in order to control for host-related factors, longitudinal samples of ethnically related patients have been obtained to study the association with clinical outcomes in HBV infection. These individuals are infected with different HBV genotypes in order to determine if an association between clinical outcome and viral genetic variability exists.

Through this course of work, knowledge discovered will aid in defining the true extent of HBV infection in the Canadian North, by determining the prevalence of occult HBV infections. In addition to this, data on the genetic variability of HBV and the association between clinical outcomes will also be obtained. This collection of data may contribute to a more 'personalized medicine' approach in HBV patients by predicting clinical outcomes, based on the genomics of the infecting HBV strain.

Hypothesis

I hypothesize the prevalence of OBI within the studied population to be 8-10%, accounting for false-positives, as a previous study determined the OBI prevalence within a Northern Canadian Inuit population to be 10% [72]. This study will also involve a similar geographically isolated population, as the previous study.

I hypothesize there will be more mutations and a higher rate of mutation within the core-coding region of the HBV genome, specifically within infected individuals of Inuit ethnicity that are associated with an adverse clinical outcome (HBV/F) in comparison to strains associated either with a benign (HBV/B6) or an unknown outcome (HBV/D).

MATERIALS AND METHODS

Patient samples

For the initial investigation into the prevalence of OBI in a Northern Canadian population, 706 archived sera were received from three First Nations settlements (numbered 1-3) in the Northwest Territories. Samples were selected from individuals in these settlements participating in a previous HBsAg serological study completed in 1985 [79]. That study determined the approximate HBsAg prevalence within the three settlements to be 11-12% [79]. All 706 archived sera were originally determined to be HBsAg negative, and 38.8% were positive for Anti-HBs. No anti-HBc testing was done during the original serosurvey. Ethics approval to perform research on these archived samples was provided by the University of Manitoba Research Ethics Committee.

The association between HBV genetic variability and clinical outcomes was studied using fifteen pairs of longitudinal samples. Five paired samples from each of the following were used: Alaska Natives infected with HBV genotype F (HBV/F), Canadian Inuit infected with HBV/B6 and Greenlandic Inuit infected with HBV/D. All paired samples were taken from the same individual at two time points, separated by 5-6 years. All samples were kept at -20˚C until analysis. The Greenland and Canadian samples were received in serum form, and the Alaskan samples were provided as extracted viral DNA, which were processed using the QIAamp® MinElute® Virus Spin Kit (Qiagen Inc, Valencia, CA). The respective laboratories provided the clinical and demographic information. Ethics approval was granted from their respective institutional boards (Alaskan Native Tribal Health Consortium [Anchorage, Alaska] and Statens Serum

Institute [Copenhagen, Denmark]) for the molecular analysis of HBV from study patients. Approval for the research on patient samples from Canada was obtained from the University of Manitoba Research Ethics Committee.

Extraction of HBV Genomic DNA

All clinical samples were extracted using silica gel filtration (QIAamp® DNA Blood Mini Kit, Qiagen Inc. Mississauga, ON, Canada) on 200 μl of sera. Briefly, 20 μl of proteinase K (20 mg/ml stock) and 200 μl of lysis buffer are added to lyse the samples via incubation at 56˚C for 10 min. Carrier RNA in the form of 10 ng yeast tRNA (10 μg/ml) was also added during this step. Following incubation, centrifugation removed drops from inside the tubes, and 230 μl of ethanol was added and the samples were vortexed. The samples were applied to the QIAamp® mini silica gel spin columns and centrifuged. Various wash and elution buffers were applied to remove contaminants and elute the DNA off the silica filter by centrifugation. From the final elution volume of 60 μl, 5 μl was used for subsequent PCR amplification and 2 μl of first stage PCR product for nested PCR reactions.

To allow for the maximum DNA isolation of low-viral load samples [80], phenol-chloroform extraction was also completed. For this method, 15 μl of proteinase K (20 mg/ml stock) and 150 μl of HBV lysis buffer (10 mM EDTA, 50 mM Tris-HCl (pH 7.5 and 1% SDS) were added to 150 μl of sera. Vortexing, and a quick centrifugation were performed, followed by incubation at 65˚C for 2.5 hrs. After lysis, 300 μl of buffer-

saturated phenol was added and the samples were inverted prior to centrifuging for 10 min at 13,000 RPM and 4˚C. The aqueous layer is added to 300 µl of a 1:1 mixture of phenol and chloroform, and again vortexed and centrifuged at the same speed, time and temperature. This was done again with the aqueous top layer into pure chloroform, and then the final aqueous phase is added into 940 µl of 100% ethanol and 150 µl of 5M-ammonium acetate. Yeast tRNA was added at 1.8 µl (10 µg/ml) for carrier RNA and the viral DNA was precipitated overnight at -20˚C. The following day, samples were centrifuged for 30 min at 4˚C and the DNA pellet washed using 70% ethanol. DNA pellets were resuspended in 25 µl. From the final volume of extract, the same volumes were used in subsequent first and nested-stage PCR amplifications as samples extracted using the Qiagen kits above.

Several precautions were taken to ensure environmental contamination of the samples did not take place. This included processing <20 samples at one time, including multiple negative extraction and amplification controls, and not processing any additional clinical samples on the same days. Finally, a "clean" workflow was also used through strict spatial and chronological workflow. These guidelines were followed through each step in the study.

Quantification of HBV viral DNA

All study samples were analyzed using the Qiagen *artus* HBV TM PCR kit (Qiagen Inc) for the detection and quantification of HBV DNA. A standard curve was produced from the following concentrations of positive controls (supplied by the Qiagen *artus* kit): 1

IU/ml, 10^1 IU/ml, 10^2 IU/ml, 10^3 IU/ml, 10^4 IU/ml and 10^5 IU/ml. The standard curve was produced from mean values, as the controls were used in duplicate. In addition, the 1 IU/ml control was tested in quadruplicate. Real-time PCR reaction mixtures for quantifying HBV consisted of 4 μl sterile water, 1 μl of supplied internal control, 15 μl HBV TM Master and 5 μl extracted viral DNA. The same mixture was employed for the standards, but with 5 μl of each positive control used. The real-time PCR (RT-PCR) was performed using the 7500 Fast Real-Time PCR System (Applied Biosystems Canada, Streetsville, ON). The FAM fluorescence was measured from each sample of HBV and the JOE fluorescence was quantified in detection of the internal control. The thermocycling conditions used were as follows: 95˚C for 10 min, followed by 45 cycles of 95˚C for 15s and 60˚C for 1 min.

Amplification of HBV Genomic DNA

The genomic DNA of HBV was amplified using the Roche Expand High Fidelity^{Plus} System, which incorporates a thermostable, proofreading DNA polymerase (Roche Diagnostics, Laval, QC, Canada). The PCR amplification contained the following in a 50 μl reaction mixture; 27.5 μl DNAse free, sterile water; 10 μl 5x reaction buffer containing 7.5 mM $MgCl_2$; 8 μl of 1.25 mM dNTPs; 1 μl each of 20 μM forward and reverse primers, 0.5 μl Expand High Fidelity^{PLUS} DNA polymerase (5 U/μl) and 2 μl DNA extract. Summarized in Table 2 are all primers used to obtain HBV sequence.

The definition of a true occult HBV infection was as follows; HBV DNA sequence confirmed in one region of the virus or PCR positive in two viral regions. Any

patients producing a positive PCR amplicon were re-amplified using primer sets for additional regions to obtain DNA sequence that was used to confirm true OBI positives.

The PhenoP1/PhenoP2 (P1/P2) primer set was utilized for the initial amplification of the full-length genome, at approximately 3.2 kb, and as described previously by Günther *et al* (1995) and Durantel *et al* (2004) [81-82]. Parameters for PCR included; initial denaturation at 94°C for 2 min, followed by 35 cycles of denaturation at 94°C for 30 seconds, annealing at 55°C for 30 seconds and extension at 68°C for 3 min. At the end of each cycle, 10 seconds were added on to the extension step for the last 25 cycles. Completion of the cycle included a final extension stage at 68°C for 10 min. The sequence obtained from this directly amplified, full-length product was assumed to be the dominant HBV species within the quasispecies of each individual.

If a positive reaction was not possible by the above-mentioned amplification, the following nested PCR reactions were performed using 2 μl of the first stage P1/P2 amplicon; “HBV-G1/PIA100”, “LLr/LLf”, “PS102/Pol2R”; “PS104/PreSpR”. These separate reactions provided “step-wise” amplification of the full-length genome. This method is summarized in Figure 5. If full-length sequence was unobtainable from these reactions, the oligonucleotides listed in Table 2 were used to amplify 2 μl of the P1/P2 first stage amplicon to provide additional sequence. All primers were used as mentioned in the corresponding reference, except HBVP-3-Diag, HBVP-4-Diag, HBVP-1-Diag, and HBVP-2-Diag. Here, the nested primers from Lindh *et al* (1998) were used as the first stage primers [83].

Figure 5. Flow chart for the full-length amplification of Hepatitis B virus isolates. This chart depicts the steps taken to ensure full-length sequence was obtained for all isolates. If the initial amplification was unsuccessful, a 'step-wise' approach was taken to obtain the remaining sequence.

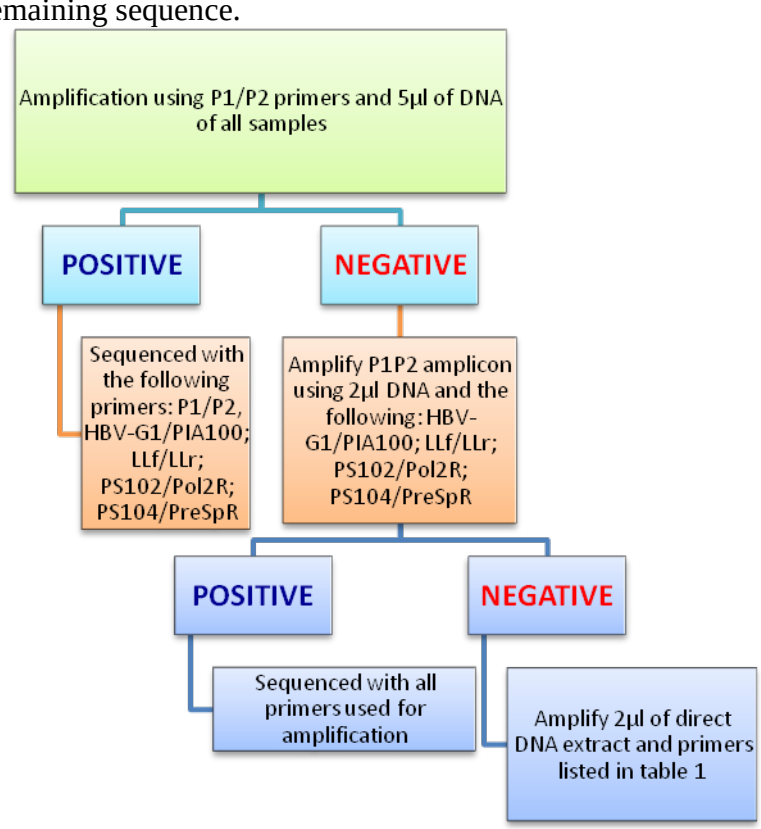


Table 2. Primers used to amplify HBV in DNA extract samples. POL – polymerase; PreC/C – Precore/core. Solid lines designate the primer sets and dotted lines separate the first and nested stages.

Name	Strand	Sequence (5'-3')	Region	Conditions	Ref.
HBVPr134	-	TGCTGCTAT GCCTCATCTTC	HBsAg	94°C 10 min-1x. 94°C 30s, 45°C	84
HBVPr135	+	CARAGACAA AAGAAAATTGG	Outer primer	30s, 72°C 30s-40x	
HBVPr75	-	CAAGGTATGTT GCCCGTTTGTC C	HBsAg	94°C 10 min-1x. 94°C 30s, 45°C	
HBVPr94	+	GGYAWAAAGG GACTCAMGATG	Nested primer	30s, 72°C 30s-35x	
HBV9POL	-	CGTCGCAGAAG ATCTCAATC	POL	94°C 10 min–1x. 94°C 30s, 50°C	85
HBV10POL	+	CCTGATGTGAT GTTCTCCATG	Outer primer	30s, 72°C 30s–40x	
HBV11nPOL	-	CCTTGGACTCA TAAGGT	POL	94°C 10 min–1x. 94°C 30s, 50°C	
HBV12nPOL	+	TTGAAGTCCCA ATCTGGATT	Nested primer	30s, 72°C 30s–35x	
HBV5PreC/C	-	GCCTTAGAGTC TCCTGAGCA	PreC/C	94°C 10 min–1x. 94°C 30s, 50°C	
HBV6PreC/C	+	GTCCAAGGAAT ACTAAC	Outer primer	30s, 72°C 30s-40x	
HBV7nPreC/C	-	CCTCACCATAC TGCACTCA	PreC/C	94°C 10 min–1x. 94°C 30s, 50°C	
HBV8nPreC/C	+	GAGGGAGTTCT TCTTCTAGG	Nested primer	30s, 72°C 30s–35x	
HBVP-3-Diag	-	CACCTCTGCCT AATCATCT	PreC/C	94°C 10 min–1x. 94°C 30s, 55°C	86, \ with mod.
HBVP-4-Diag	+	TGACTACTARD TCCCTGGATGC T	Outer primer	30s, 72°C 30s–40x	
HBVP-1-Diag	-	GCTGTGCCTTG GGTGGCTTT	PreC/C	94°C 10 min–1x. 94°C 30s, 55°C	
HBVP-2-Diag	+	GTATGGTGAGG TGAACAATG	Nested primer	30s, 72°C 30s–35x	

Table 2 cont. Primers used to amplify HBV in DNA extract samples. POL – polymerase; PreC/C – Precore/core. Solid lines designate the primer sets and dotted lines separate the first and nested stages.

Name	Strand	Sequence (5'-3')	Region	Conditions	Ref.
Pheno P1	+	CCGGAAAGCTTATGCTC TTCTTTTTCACCTCTGCC TAATCATC	Full length	94˚C, 2 min-1x. 94˚C, 30s; 55˚C, 30s; 68˚C, 3 min – 10x. 94˚C, 30s; 55˚C, 3 min +10s/cycle- 25x	82
Pheno P2	-	CCGGAGAGCTCATGCTC TTCAAAAAGTTGCATGG TGCTGGTG	Outer primer		
HBV-G1	+	TCACCATATTCTTGGGA ACAAGA	PreS	94˚C, 2 min-1x. 94˚C, 30s; 50˚C, 30s; 68˚C, 40s–35x	83
HBV-G2	-	TTCCTGAACTGGAGCCA CCA	Outer primer		
C2F	+	ATGGMRACCACCGTGA ACGC	Core	94˚C, 2 min-1x. 94˚C, 30s; 50˚C, 30s; 68˚C, 40s–35x	In-house
C2R	-	YCCCACCTTATGWGT CCAAGG	Outer primer		
PreS1 F	+	AGGTRGGAGYGGGAGC ATTCGG	PreS	94˚C, 2 min– 1x. 94˚C, 30s; 55˚C, 30s; 68˚C, 40s –35x	In-house/ 83
HBV-G2	-	TTCCTGAACTGGAGCCA CCA	Outer primer		
PS104	+	See below	Pol/PreS	94˚C, 2 min– 1x. 94˚C, 30s; 45˚C, 30s; 68˚C, 40s – 35x	In-house
PreSp R	-	CCCAAGCTTCTTAGAGG TGGAGAGAGGG	Outer primer		
EP1	+	See below	Core	94˚C, 2 min– 1x. 94˚C, 30s; 55˚C, 30s; 68˚C, 40s – 40x	In house
C2R	-	YCCCACCTTATGWGTCC AAGG	Outer primer		
C2F	+	ATGGMRACCACCGTGA ACGC	Core primer	94˚C, 2 min– 1x. 94˚C, 30s; 50˚C, 30s; 68˚C, 40s – 35x	
survC R	-	GACGCGGYGATTGAG	Nested primer		
EP1	+	GCATGGAGACCACC GTGAAC	Core	94˚C, 2 min– 1x. 94˚C, 30s; 55˚C, 30s; 68˚C, 40s – 40x	87
EP2	-	GGAAAGAAGTCAGAAG GCAA	Outer primer		
EP3	+	CATAAGAGGACTCTTGG ACT	Core	94˚C, 2 min-1x. 94˚C, 30s; 55˚C, 30s; 68˚C, 40s – 35x	
EP4	-	GGCAAAAAAGAGAGTA ACTC	Nested primer		

Table 2 cont. Primers used to amplify HBV in DNA extract samples. POL – polymerase; PreC/C – Precore/core; ENHI/X-Enhancer I/HBx protein region. Solid lines designate the primer sets and dotted lines separate the first and nested stages.

Name	Strand	Sequence (5'-3')	Region	Conditions	Ref.
sG145Rf	+	TGGAGAACATCGCATCAGG	Surface	94˚C, 2 min - 1x. 94˚C, 30s; 55˚C,	88
sG145Rr	-	TGGTAACAGCGGCATAAAGG	Outer primer	30s; 68˚C, 40s - 35x	
Spr1A	+	GTTCAGGAACAGTAAGCCC	Surface	94˚C, 2 min – 1x. 94˚C, 30s; 56˚C,	89
DRv2as	-	AGAAAGGCCTTGTAAGTTGGCGA	Outer primer	30s; 68˚C, 40s – 40x	
Spr2A	+	ACTTTCCAATCAATAGGCC	Surface	94˚C, 2 min – 1x. 94˚C, 30s; 56˚C,	
DRv2s	-	CGTGGTGGACTTCTCTCAATTTTC	Nested primer	30s; 68˚C, 40s – 35x	
PolF1	+	CYTTYGGAGTGTGGATTCGC	POL	94˚C, 2 min – 1x. 94˚C, 30s; 55˚C,	16
PolR1	-	TGGGATGGGAATACARGTGC	Outer primer	30s; 68˚C, 40s – 40x	
PolF2	+	CGTTTGTCCTCTAMTTCCAGG	POL	94˚C, 2 min – 1x. 94˚C, 30s; 55˚C,	16
PolR2	-	ACGTARACAAAGGACGTCCC	Outer primer	30s; 68˚C, 40s – 35x	
LLf	+	TCCTGCTGGTGGCTCCAG	Surface	94˚C, 2 min – 1x.	In-house
LLr	-	CGTTGACATACTTTCCAATCAA	Outer primer	94˚C, 30s; 55˚C, 30s; 68˚C, 40s – 40x	
nLLf	+	ACCCTGYRCCGAACATGGA	Surface primer	94˚C, 2 min – 1x. 94˚C, 30s; 55˚C,	
nLLr	-	CAACTCCCAATTACATARCCCA	Nested primer	30s; 68˚C, 40s – 35x	
Sybr51	+	GCTGACGCAACCCCCACT	ENHI/X	94˚C, 2 min – 1x.	90
Sybr52	-	AGGAGTTCCGCAGTATGG	Outer primer	94˚C, 30s; 55˚C, 30s; 68˚C, 40s – 35x	
HBx1093F	+	ACTTTCTCGCCAACTTACAAGGCCTTT	HBx	94˚C, 2 min – 1x. 94˚C, 30s; 55˚C,	91
HBx1977R	-	GAAGGAAAGAAGTCAGAAGGCAA	Outer primer	30s; 68˚C, 40s – 40x	
HBx1265F	+	GCCGATCCATACTGCGGAACTC	HBx	94˚C, 2 min – 1x. 94˚C, 30s; 55˚C,	
HBx1887R	-	GGCAAAAAAGAGAGTAACTC	Nested primer	30s; 68˚C, 40s – 35x	

Table 2 cont. Primers used to amplify HBV in DNA extract samples. POL – polymerase; PreC/C – Precore/core; ENHI/X-Enhancer I/HBx protein region. Solid lines designate the primer sets, and dotted lines separate the first and nested stages.

Name	Strand	Sequence (5'-3')	Region	Conditions	Ref.
PS102	+	CCTATTGATTGGAAAGTATGTCAA	EHNI/X/PreC	94˚C, 2 min – 1x. 94˚C, 30s; 56˚C, 30s; 68˚C, 40s- 40x	
PA7	-	TATGCCTACAGCCTCCTAATACAA	Outer primer		
PIS102	+	CGTATTGTGGATCCTTTGGGTTT	EHNI/X/PreC	94˚C, 2 min – 1x. 94˚C, 30s; 56˚C, 30s; 68˚C, 40s – 35x	
PIA7	-	TCAAGCTTCTCCCAGTCTTTAAAC	Nested primer		
PS103	+	GCCCAATGTCTTACATAAGAGGAC	Core	94˚C, 2 min – 1x. 94˚C, 30s; 56˚C, 30s; 68˚C, 40s – 40x	
PA104	-	AAAGTTTCCCACCTTATGAGTCCA	Outer primer		
PIS103	+	CTTGGATCCTCTGTAATGTCA	Core	94˚C, 2 min – 1x. 94˚C, 30s; 56˚C, 30s; 68˚C, 40s – 35x	92
PIA104	-	CCAAGCTTTACTAACATTGAG	Nested primer		
PS104	+	CCCTATCTTATCAACACTTCC	POL/PreS	94˚C, 2 min – 1x. 94˚C, 30s; 56˚C, 30s; 68˚C, 40s – 40x	
PA100	-	GAAGTCCACCACGAGTCTAGA	Outer primer		
PIS104	+	TTCCGGATCCTACTGTTGTTA	POL/PreS	94˚C, 2 min – 1x. 94˚C, 30s; 56˚C, 30s; 68˚C, 40s – 35x	
PIA100	-	GGTATTGTGAGGAAGCTTGTC	Nested primer		
PS100	+	ATCCTCAGGCCATGCAGT	PreS/Surface	94˚C, 2 min – 1x. 94˚C, 30s; 56˚C, 30s; 68˚C, 40s – 40x	
PA102	-	CGTCAGCGAACACTTGG	Outer primer		
PIS100	+	CCACCAAACTCTTCAGGATCC	PreS/Surface	94˚C, 2 min – 1x. 94˚C, 30s; 56˚C, 30s; 68˚C, 40s – 35x	
PIA102	-	AACGGGGTAAAGCTTCAGATA	Nested primer		
Xf07	+	CTCCTCTGCCGATCCATACTGCGGAACTCC	HBx/PreC	94˚C, 2 min–1x. 94˚C, 30s; 50˚C, 30s; 68˚C, 40s– 35x	In-house
EP2	-	See above	Outer primer		
HBV-G1	+	See above	PreS	94˚C, 2 min – 1x. 94˚C, 30s; 50˚C, 30s; 68˚C, 40s – 35x	
PIA100	-		Nested primer		

HBV DNA Amplicon Gel Purification

All PCR amplicons and cloning products were analyzed by gel electrophoresis, using a 1.5% agarose gel (UltraPure™ Agarose, Invitrogen Inc.) in 1X TAE (Bio-Rad Laboratories) with 5 μl SYBR® Safe DNA gel stain added for visualization under UV light (Invitrogen Inc.). Each gel was run at 90V for 60 min and positive PCR amplicons were gel-purified using the Qiagen MinElute® gel extraction kit. The appropriate DNA fragment was excised from the gel using sterile blades and placed into sterile 1.5ml microtubes. The slices were weighed and three volumes of solubilization buffer to one volume gel slice were added, and incubated for 10 min (or until dissolved) at 50˚C with shaking. Isopropanol was added at 1 gel volume and inverted. The mixture was then applied to the silica gel column for binding and washing and a final eluate of 10 μl was obtained.

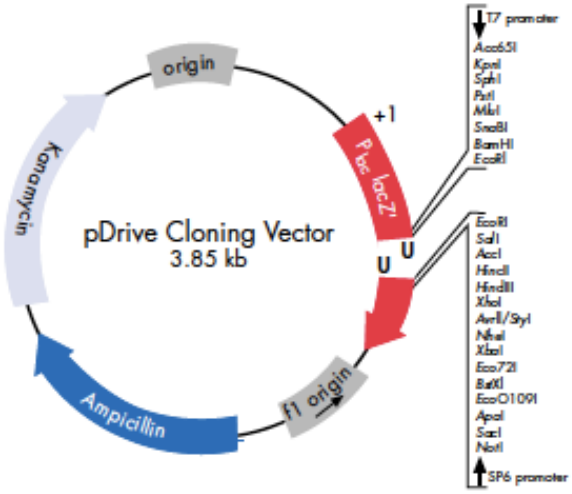
Cloning of PCR Products

Specific regions of the fifteen paired samples were cloned to investigate the putative quasispecies nature of each strain. To investigate the PreC and core regions, the primer set EP1/C2R were used in the amplification, followed by nesting with either: C2F/survCR or EP3/survCR. Sequence of the HBsAg coding region was amplified using the LL/nLL primer sets. For ligation of the amplicon, 4 μl of the purified DNA was added to 5 μl ligation master mix (2x) and 1 μl of the pDrive cloning vector (50 ng/μl) (Qiagen Inc.). The ligation was mixed vigorously for 25s and incubated for 2 hrs at 4˚C. If cloning was continued at a later date, the ligation mixture was stored at -20˚C. The ligation was transformed into 50 μl One Shot® TOP10 chemically competent *E.coli* cells

(Invitrogen Canada Inc, Burlington, ON) using 5 μl of the ligation mixture. The cells were kept at -80˚C until use, at which point they were thawed on ice to prevent any loss of viability. The mixture was incubated for 30 min on ice, followed by heat shocking for 30 sec at 42˚C. A pre-warmed mixture of super optimal broth (250 μl) with catabolite repression (SOC) media (Invitrogen Canada Inc.) was added to the transformation mixture and incubated with shaking for 1 hr at 37˚C and 225 rpm. For blue-white screening of the clones, X-gal (50 μg/ml) was added to ampicillin (100 μg/ml)-containing lysogeny broth (LB) plates, and was allowed to diffuse into the plates for 30 min at room temperature and then placed at 37˚C for 30 min. As shown in Fig. 6, ampicillin was the antibiotic of choice as our vector contained resistance. The plates were brought to room temperature from 4˚C for at least 1 hr prior to plating. Approximately 150 μl of the transformation mixture was added to two plates and incubated at 37˚C overnight.

On average, 30 clones per sample were chosen from the plates and a minimum of 20 were used in the analysis, to sample the quasispecies population. Each separate colony was grown in 3 ml of LB-ampicillin (100 μg/ml) broth overnight at 37˚C with shaking. The following day, cells were harvested by transferring the cultures into 1.5 ml microtubes and centrifuging at 13,000 rpm for 2-3 min at room temperature. Purification of the plasmid DNA was done using QIAprep® Miniprep kits (Qiagen Inc.) either manually or semi-automatically using a QIAcube (Qiagen Inc.). To begin, the pelleted bacterial cells were resuspended into buffer, lysed and neutralized during a 10 min centrifugation step. The supernatants were applied to a spin column and the plasmid DNA washed according to the manufacturer's instructions. Using the provided buffer, a final eluate of 50 μl was prepared.

Figure 6. pDRIVE (Qiagen, Inc) cloning vector map. This vector was used in the clonal analysis aspect to investigate the putative quasispecies nature of strains.



Each clone was then digested with the restriction enzyme, EcoR1 (New England Biolabs, Pickering, ON) for 2 hrs at 37˚C. The digested samples were analyzed using gel electrophoresis to detect a specific insert. A 1.5% agarose gel (UltraPure™ Agarose, Invitrogen Inc.) made with 1X TAE (Bio-Rad Laboratories, Mississauga, ON) was prepared for the analysis. SYBR® Safe DNA gel stain (10,000x; Invitrogen Inc.) was added (6 µl) for visualization under UV light. Each gel was run at 90V for 60 min and any clones carrying a putative HBV sequence were detected by the presence of the digested band corresponding to the size of original amplicon. Some samples could not be included in this aspect of the study, as they did not produce a successful cloning reaction.

HBV DNA Sequencing

All purified positive PCR amplicons were sequenced by the Genomics Core Facility

within the National Microbiology Laboratory (NML) in Winnipeg, MB. Sequencing was performed using an ABI Prism 3100 Genetic Analyzer (Applied Biosystems, Streetsville, ON, Canada) and BigDye v3.1 Terminator Chemistry. Positive DNA amplicons were sequenced with the primers used for amplification. Positive clonal samples were sequenced using the M13 primers (M13 -20 Forward: 5'GTAAAACGACGGCCAGT-3'; M13 Reverse: 5'AACAGCTATGACCATG-3').

HBV DNA Sequence Analysis

HBV DNA sequences were analyzed using the Lasergene sequence analysis software package (v7.1.0; DNAStar Inc., Madison, WI). The genotype of each isolate was initially evaluated using the NCBI HBV Genotyping Tool [93] and later confirmed by phylogenetic analysis through alignment with five genotype representatives sequences (described below) using ClustalW (ClustalX v1.81) [94] and MEGA v4 [2].

The most appropriate substitution model for analysis of full genome sequence data was provided by FindModel (http://www.hiv.lanl.gov/content/sequence/findmodel/findmodel.html) which determined the most appropriate substitution model to be General Time Reversible (GTR) plus Gamma, using a gamma parameter of 0.02. A phylogenetic tree of all full-length longitudinal samples was produced using DIVEIN© online software [1] with 100 bootstrap replicates and the BioNJ algorithm. Tree topology and branch length optimization was performed using nearest neighbour interchanges (NNIs), subtree pruning and regrafting (SPR) algorithms. A tree was also constructed for solely the genotype D isolates using the methods mentioned above.

The GenBank accession numbers of the representative sequences used in the

determination of HBV genotype and subgenotype were AF297625, AY233279, AB116087 (HBV/A1); AF090841, AY152726, AB116076 (HBV/A2); D50522, D23678 (HBV/B1); AF121245 (HBV/B2); M54923 (HBV/B3); AB031266 (HBV/B4); AB219428 (HBV/B5); AB287315, AB287321, DQ463799 (HBV/B6); AB033556 (HBV/C1); AF223954 (HBV/C2); AB104709, AF280817 (HBV/D1); AY090453, AB078032 (HBV/D2); X85254, AJ344117, DQ315776 (HBV/D3); AB033559, AB048701, AB048703, AB048702 (HBV/D4); X75664 (HBV/E); AY179735 (HBV/F1); AY311369 (HBV/F2); AB036915 (HBV/F3); X75658 (HBV/F4); AF160501 (HBV/G) and AY090460 (HBV/H).

Nucleotide pairwise distances of full-length genomes among the different genotypes were analyzed by DIVEIN© using the Tamura - Nei (TN93) model. Further nucleotide distance and genetic diversity analysis was performed using the following substitution models, as determined by FindModel, to be appropriate for individual genotype coding region alignments: Felsenstein 1981 (F81); Jukes - Cantor 1969 (JC69); Hasegawa, Kishino and Yano 1985 (HKY85); GTR and Tamura - Nei (TN93). The DIVEIN® parameters included: optimized equilibrium frequencies, estimated proportion of invariable sites, and four substitution rate categories. The gamma (γ) distribution parameter was either uniform or equal to 0.04, depending on the suggestion of FindModel. For the calculation of amino acid distances, MEGA4 software was used with the following parameters; pairwise deletion for gaps and missing data, and p-distance as the amino acid model, with all other parameters set to default. For the analysis of non-overlapping regions, AA49-180 (nt 2456-2853) is non-overlapping region "POL1" and AA582-760 (nt 833-1373) is "POL2".

Phylogenetic analysis of the HBV DNA sequence from OBI-positive individuals was performed with some modifications to that mentioned above. Using DIVEIN® online software, the Kimura 2-parameter model was used along with the neighbour-joining model for inferring the tree. Subgenotyping of the OBI-positive patients was performed using phylogenetics.

Quasispecies Analysis

Clonal sequences were aligned using ClustalW (ClusterX v1.81) and analyzed subsequently using MEGA4 software [2]. The genetic distance at the nucleotide level within the non-overlapping portion of the HBcAg was calculated using Tamura 3-parameter model with transitional and transversion rates and G+C content bias. The genetic diversity was calculated at the AA level using the p-distance model. In determination of the dS and dN levels of all quasispecies, the modified Nei-Gojobori model with Jukes-Cantor correction was employed.

Statistical Analyses

In the analysis of OBI prevalence within Northern Canadian communities, logistical regression was performed using OBI-positives as the dependant variable. One-way analysis of variance (ANOVA) was performed to examine the differences throughout the sequence data. Pearson's test of correlation was employed to determine if a relationship between nucleotide distance and HBV viral load existed amongst genotypes B, D and F. *P* values <0.05 were considered significant. Mean values are stated, followed by the standard error (SE). All statistical analyses were performed using SPSS v.16 (Chicago, Illinois).

RESULTS

1. Patient samples

1.1 Occult-HBV patients

We studied 706 HBsAg-negative patients in the determination of OBI within a Northern Canadian population. Almost all individuals in the studied population (98.2%) were of First Nations descent, with the remaining 1.8% either Caucasian or Asian. Of the three settlements, more than half of the patients resided in settlement 1. The total population was 51.9% female with the median and mean age of patients at 15.57 and 22.21, respectively. Based on the specified definition (HBV DNA sequence confirmed in one genomic region or PCR positive in two regions), the prevalence of OBI within the studied population is 3.8% (27/706). Within the 27 OBI-positive cases, almost twice as many females (62.9%) were positive than males. Of all anti-HBs-positive patients, 4% (11/274) were also occult-HBV positive. As expected, higher numbers of OBI-positive patients resided in settlement 1 ($p<0.05$). Most patients involved were of First Nations descent and the distribution of age was similar in OBI-positive and -negative patients.

Table 3. Demographic and serological information for occult hepatitis B virus (HBV)-positive and -negative individuals residing within three settlements in Northern Canada. *$p<0.05$.

	Occult-HBV Positive, *n* (%)	**Occult-HBV Negative, *n* (%)**	***p***
Settlement 1	22 (81.5)	424 (62.5)	0.011*
Female	17 (62.9)	349 (51.5)	0.241
First Nations	25 (92.5)	666 (98.2)	0.039*
Age (0-9 years)	8 (29.6)	164 (24.2)	0.668
Anti-HBs positive	11 (40.7)	263 (38.8)	0.838
Total	27 (100)	678 (100)	

1.2 HBV Genotypic comparison patients

The patients investigated for the association between clinical outcome and genetic variability are listed in Table 4 along with their clinical and demographic information. One third of the patients studied were HBV/B-infected Canadian Inuit with no clinical or biochemical evidence of active or advanced liver disease during the period of study. All HBV/B patients had HBeAg seroconverted (HBe-SC) prior to the study period and had an average age of 66 years at baseline. Patient samples were also collected from Alaska Natives infected with HBV/F, all of whom developed HCC within 2 years of the second time point. Alaskan patients had an average age of 30.6 years at the first time point. Three HBV/F patients had HBe-SC during the study period (patients 1, 2 and 5), with one having previous history of HBeAg sero-reversion (patient 2). Alaskan patient 3 remained HBeAg positive during the study, with patient 4 seroconverting prior to the study. The remaining third of studied patients came from a Greenlandic Inuit population infected with HBV/D, with an average age of 33.2 years at the initial time point. Although the clinical outcome of the HBV/D-infected individuals is currently unknown in this population, this genotype has been associated with varying degrees of liver disease in other populations. Three of the HBV/D-infected individuals had HBe-SC prior to the study (pairs 3-5) with the remaining two patients HBeAg positive throughout the study period. Patients infected with each of the three implicated genotypes display highly variable outcomes.

Table 4. Clinical and demographic information for hepatitis B virus (HBV)-infected circumpolar patients involved in the association between clinical outcome and genetic variability of HBV study. Viral load measurements were determined using the *artus* HBV TM PCR kit (Qiagen Inc.). All other clinical and demographic information came from the respective laboratory. HBeAg= HBV e antigen.

Patient #	Sample year	Sex	Age at sample	Viral load (Log_{10} IU/ml)	HBeAg
			Greenland HBV/D		
1a	1998	F	8	8.64	+
1b	2004	F	14	4.97	+
2a	1998	F	41	3.62	+
2b	2004	F	47	3.48	+
3a	1998	M	42	2.96	-
3b	2004	M	48	2.85	-
4a	1998	M	35	3.49	-
4b	2004	M	41	2.71	-
5a	1998	F	40	2.97	-
5b	2004	F	46	2.65	-
			Mean	**3.83**	
			Alaskan HBV/F		
1a	1975	M	58	11.70	+
1b	1980	M	63	3.56	-
2a	1982	F	14	8.85	+
2b	1988	F	20	7.45	-
3a	1988	M	45	6.94	+
3b	1992	M	49	4.80	+
4a	1987	M	11	3.49	-
4b	1992	M	16	3.58	-
5a	1988	M	25	6.30	+
5b	1993	M	30	3.58	-
			Mean	**6.02**	
			Canadian HBV/B		
1a	2004	M	54	2.24	-
1b	2009	M	59	3.47	-
2a	2004	M	62	2.00	-
2b	2009	M	67	3.20	-
3a	2004	M	63	2.95	-
3b	2009	M	68	2.12	-
4a	2004	M	66	3.58	-
4b	2009	M	71	4.24	-
5a	2004	F	85	2.60	-
5b	2009	F	90	3.16	-
			Mean	**2.96**	

2. Viral Load Analysis of HBV-infected patients

The viral load of all OBI-positive samples was measured, but only two were quantifiable, at 2.86 Log_{10} IU/ml and 2.40 Log_{10} IU/ml for occult 4 and 8, respectively. To determine the association between clinical outcome and HBV genetic variation, the average viral load for the three genotypic groups was; Alaska HBV/F: 6.02 Log_{10} IU/ml; Greenland HBV/D: 3.83 Log_{10} IU/ml and Canada HBV/B: 2.96 Log_{10} IU/ml. There was a significant difference between the viral load levels of genotype B and F and between genotype D and F ($p<0.05$). The viral loads decreased for the Alaskan patients over the 5-year period, but remained comparatively high to the other two genotypes. On average, the Canadian HBV/B patients had a slight increase in viral load over time, whereas the Greenland viral DNA levels declined over the sampling period. The Alaskan HBV/F isolates had the highest average viral load and the Canadian HBV/B-infected samples had the lowest, on average.

3. Phylogenetic Analysis

3.1 Occult-HBV phylogeny

Genotyping of the OBI positive genomes displayed the following distribution; genotype A: 48.1% (13/27); genotype B: 33.3% (9/27) and genotype D: 18.5% (5/27). All OBI-positive patients, except one (Occult 8) were included in either or both of the phylogenetic trees in Fig. 7, based on the sequence available. Occult 8 was genotyped as HBV/D using the National Center for Biotechnology Institute (NCBI) viral genotyping tool (http://www.ncbi.nlm.nih.gov/projects/genotyping/formpage.cgi), due to the short

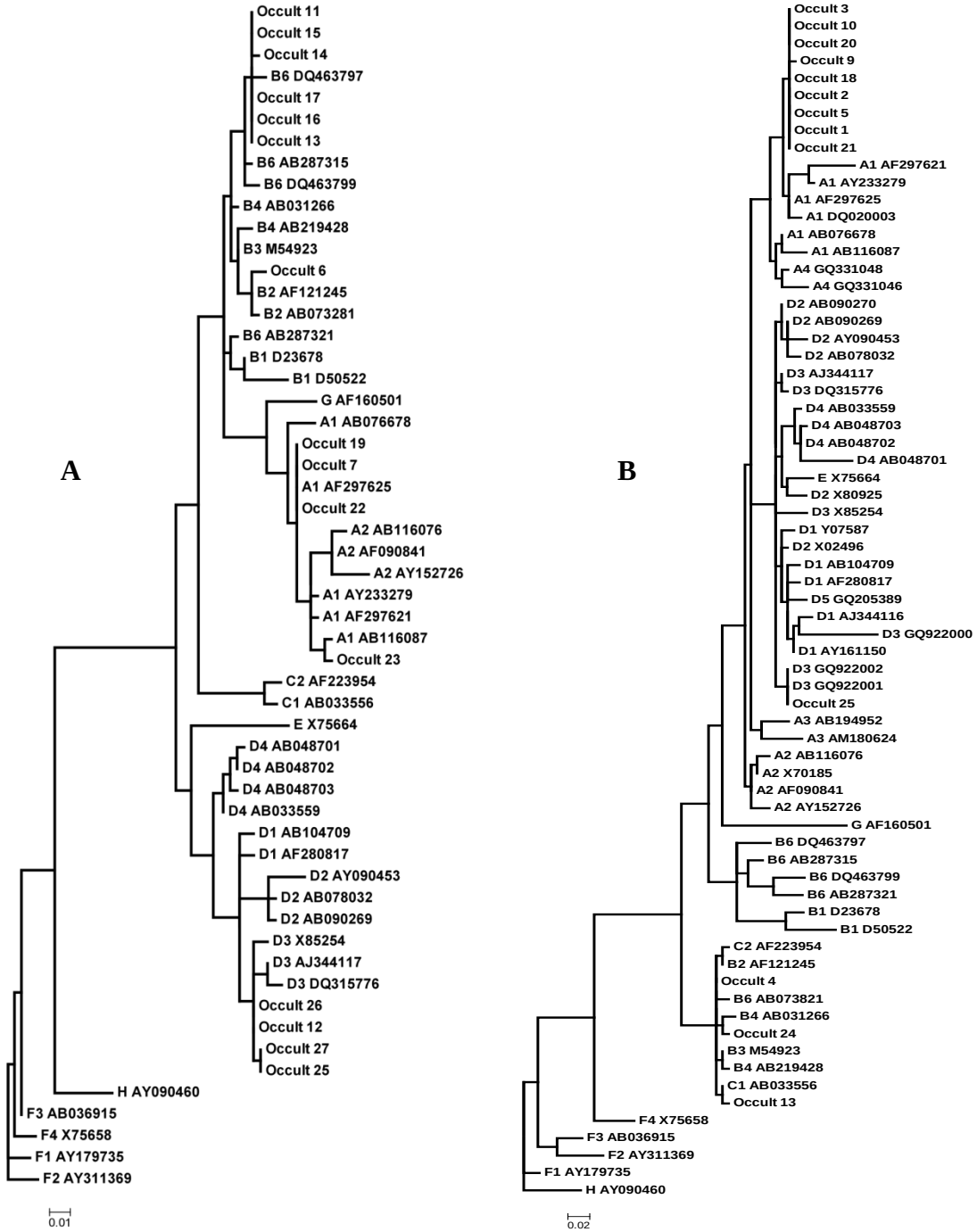


Figure 7. Phylogenetic analysis of Northern Canadian occult-hepatitis B virus positive patients based on partial surface (HBsAg) (313 bp) (**A**) or core (HBcAg) (177 bp) (**B**) sequence. The tree is constructed using DIVEIN© online software [1]. Comparative GenBank sequences are designated by the HBV genotype or subgenotype, followed by the accession number. The ruler shows the branch length for pairwise distances equal to either 0.01 (**A**) or 0.02 (**B**).

sequence available for this sample. Occult-positive samples are currently being re-extracted and amplified to further confirm true OBI positives. The majority of the occult HBV-infected samples were genotype A (48.1%).

3.2 HBV genotypic comparison phylogeny

A phylogenetic tree was constructed using full-length HBV sequences from fifteen longitudinal paired samples using DIVEIN© online software [1] (Fig. 8). Although they cluster together, separate from each subgenotype reference clade, the HBV/D-infected patients share approximately 97.4-97.7% identity with subgenotypes D2 and D1, respectively. Greenland pair 5 did show >4% divergence with D1 and D2, suggesting a possible unique, new subgenotype. Subgenotyping of the Greenland patients was further confirmed using multiple HBV/D subgenotype references, as shown in Fig. 9. The Alaskan patients have the highest homology to subgenotype F1 references, with the Canadian patients most closely related to HBV/B6.

Collectively, the Canadian HBV/B6 genomes are more highly branched and are further from their putative ancestor than that of the other genotypes from their own ancestors, indicating higher divergence for B6 (Fig. 8). The Alaskan HBV/F genomes show comparatively very little divergence from each other and over time, compared to HBV/D- and HBV/B6-infected patients. Here, phylogenetic analysis has helped to confirm the genotype, as well as provide preliminary nucleotide distance relationships from a putative ancestor for the three genotype groups under study.

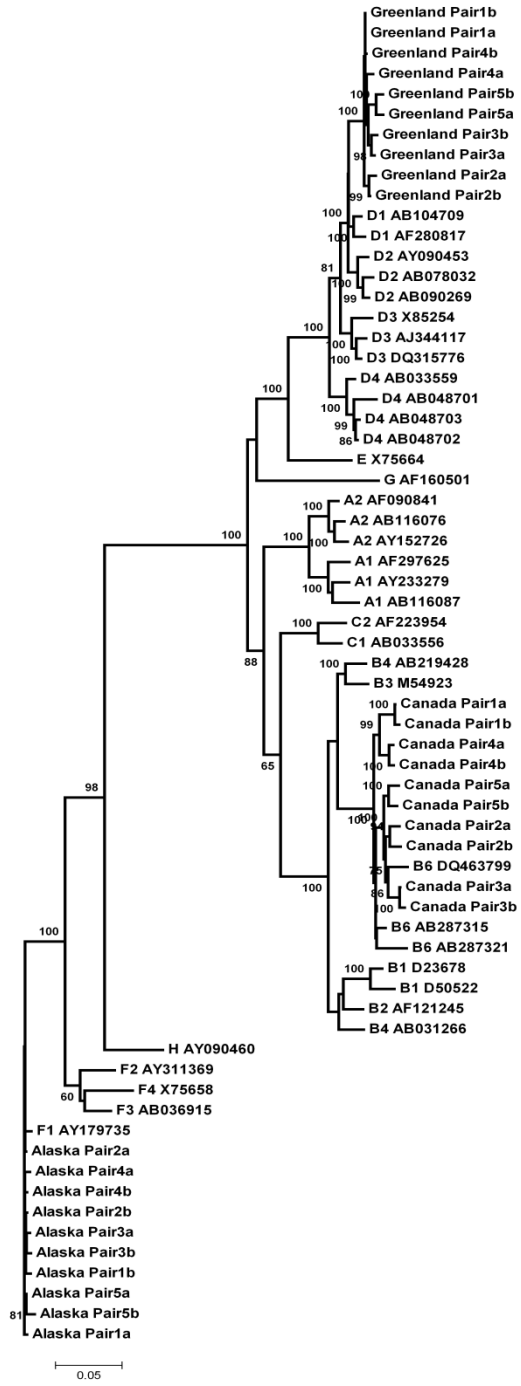


Figure 8. Phylogenetic tree of 15 longitudinal full-length circumpolar hepatitis B virus (HBV) isolate pairs. Sample codes are described in Table 4 and the tree was constructed using DIVEIN© online software [1]. Comparative GenBank sequences are designated by the HBV genotype or subgenotype, followed by the accession number. Bootstrap confidence values of ≥60% are given. The ruler shows the branch length for a pairwise distance equal to 0.05

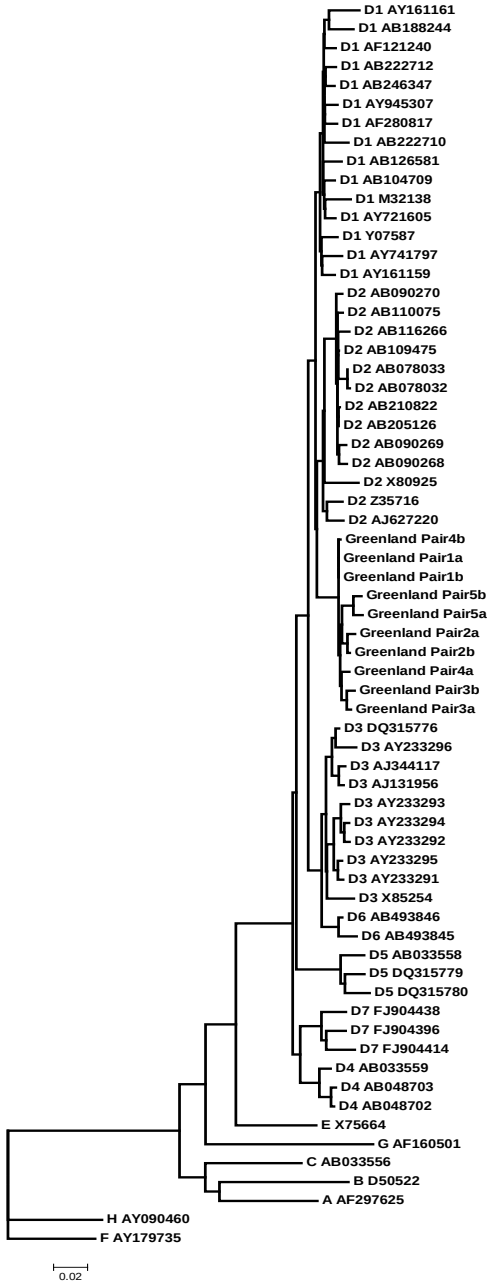


Figure 9. Phylogenetic tree of five longitudinal full-length hepatitis B virus (HBV)-genotype D isolate pairs from Greenland. Sample codes are described in Table 4 and the tree was constructed using DIVEIN© online software [1]. Comparative GenBank sequences are designated by the HBV genotype or subgenotype, followed by the accession number. Bootstrap confidence values of ≥60% are given. The ruler shows the branch length for a pairwise distance equal to 0.02.

4. Sequence Analysis of Full-Length Genomes

The Greenlandic and Canadian HBV DNA were initially extracted using commercial kits, but as many Greenlandic patient samples were continually negative for various PCR reactions, these serum samples were also extracted using phenol-chloroform. This method allowed for many successful amplification reactions providing full-length sequence. Only two Alaskan patient samples could be successfully amplified for a full-length product using the P1P2 primer set, followed by subsequent sequencing using primers from Table 2. Most samples required further amplification reactions on the P1P2 product in order to achieve full-length sequence, via "step-wise" amplifications.

The full-length sequence of fifteen longitudinal patient pairs infected with HBV, inhabiting the circumpolar regions of the world was determined. All genotype B and F isolates were 3215 nt in length, with the exception of HBV/F pairs 1 and 2, which were 3206 nt in length, owing to deletions. All genotype D patients were 3182 nt in length, with the exception of HBV/D patient 2a, which was 3185 nt owing to a 3-bp insertion. The sequence of the direct amplicon was assumed to be the dominant HBV species present within the quasispecies population.

Figure 10 lists the nucleotide substitutions observed within the fifteen longitudinal patient paired samples following sequencing. Substitutions that possess an association with or an independent risk factor for cirrhosis and/or HCC in genotypes B and/or C were investigated and listed, although other mutations such as insertions/deletions are also listed. Although these mutations were detected in all three genotypes, deletions that are commonly associated with chronic hepatitis and HCC cases were found only in HBV/F-infected patients. Insertions have also been associated with adversity, and HBV/D pair 2

displayed a single AA insertion within the non-overlapping portion of the X region. The common double substitution at nt1762/1764, along with mutations C1766T and T1768A, are believed to be associated with advanced liver disease and interestingly, were found in the two HBV/F pairs also displaying the highest viral loads of all patients. All HBV/B6-infected patients and most of the HBV/D patients lacked these multiple BCP substitutions, yet were found to display the mutation known to abolish production of the HBeAg, G1896A. Although Greenland pair 2 was HBeAg positive at the first time point, the 1896A mutation was also found. This patient may have recently obtained this mutation; hence there may have been residual HBeAg present when testing was completed in 1998. As expected, 1896A was detected more often in HBV/D, than in HBV/F [18]. In contrast to the HBV/B6 patients, those infected with HBV/F were more commonly found to acquire the 1896A mutation over time, and HBV/D patients lost the mutation.

Overall, the majority of substitutions within all 3 genotypes could be found in the BCP/PreC/Core regions. HBV/F patients have more mutations associated with adverse outcomes within the X/BCP regions than the other two genotypes, whereas substitutions were concentrated within the PreC/Core regions for HBV/D and HBV/B6 patients. The only region to display sustained mutants amongst all 3 genotypes was within the ENHI region. No patients possessed known substitutions within the non-overlapping portion of the POL; in addition, the B6- and D-infected patients had no mutations within the overlapping HBsAg. Certain pairs, such as HBV/D pair 5, were noted to display more mutations than other pairs. This pair had a mutation profile similar to that of HBV/F pairs, and was found to possess mutations not found in any other HBV/D or B6 patients.

Of the 22 substitutions evaluated, genotypes B6 and D were noted to maintain many of these mutations over time, whereas HBV/F patients made more acquisitions during the sampling period. These findings assist in demonstrating that the presence of specific mutations is associated with certain clinical outcomes.

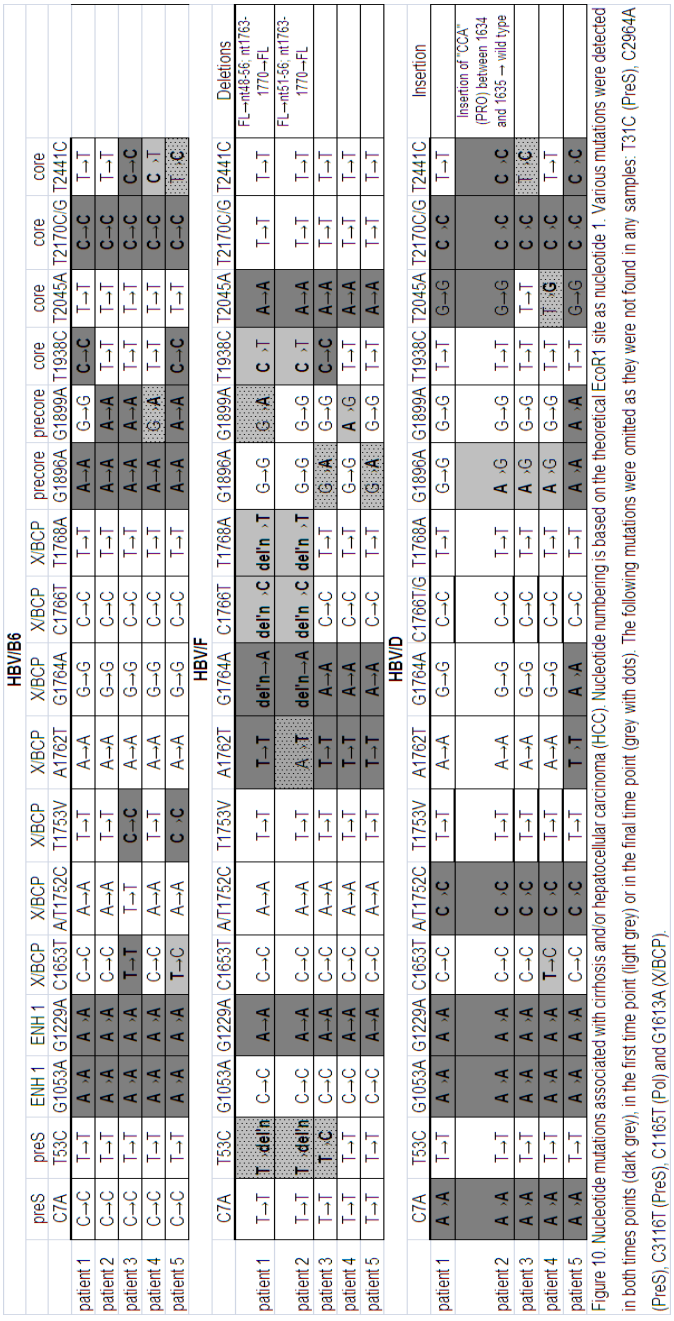


Figure 10. Nucleotide mutations associated with cirrhosis and/or hepatocellular carcinoma (HCC). Nucleotide numbering is based on the theoretical EcoR1 site as nucleotide 1. Various mutations were detected in both times points (dark grey), in the first time point (light grey) or in the final time point (grey with dots). The following mutations were omitted as they were not found in any samples: T31C (PreS), C2964A (PreS), C3116T (PreS), C1165T (Pol) and G1613A (X/BCP).

5. Genetic distance and diversity amongst genotypes

The overall mean nucleotide distance between and within the three genotypic groups based on the full-length sequence is shown in Fig. 11. The pairs of HBV/B6 have a significantly higher mean nucleotide distance at 4.1% in comparison to HBV/D and HBV/F pairs, at 1.3% and 0.7%, respectively (p<0.0001). The overall mean distance within each pair of the three genotypes was insignificant. The genotype with the benign outcome, HBV/B6, is determined to have the highest genetic distance, when compared to the genotype associated with HCC (HBV/F).

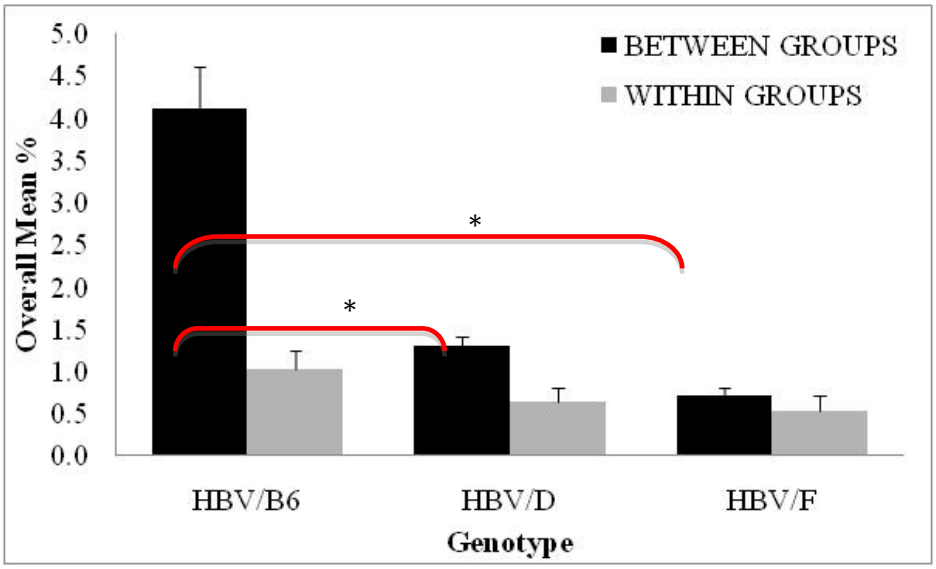


Figure 11. Overall mean nucleotide distance between and within individual genotypic groups based on full-length hepatitis B virus (HBV) sequence. Pairwise distances were calculated using the Tamura-Nei model ($\gamma = 0.04$) in MEGA4 software [2]. *p<0.0001. Standard error bars are shown.

The nucleotide distance of only non-overlapping regions was examined to control for the evolutionary constraints posed on the overlapping portions of the HBV genome (Fig. 12). The HBV/B6 genomes display the largest distance, followed by HBV/D- and F-infected patients throughout all five regions analyzed and concurs with the findings shown by the full-length diversity. The nucleotide distance within the PreC/C region was almost three times greater than any other non-overlapping region and was highest amongst the HBV/B6-infected patients (1.87%±1.45%).

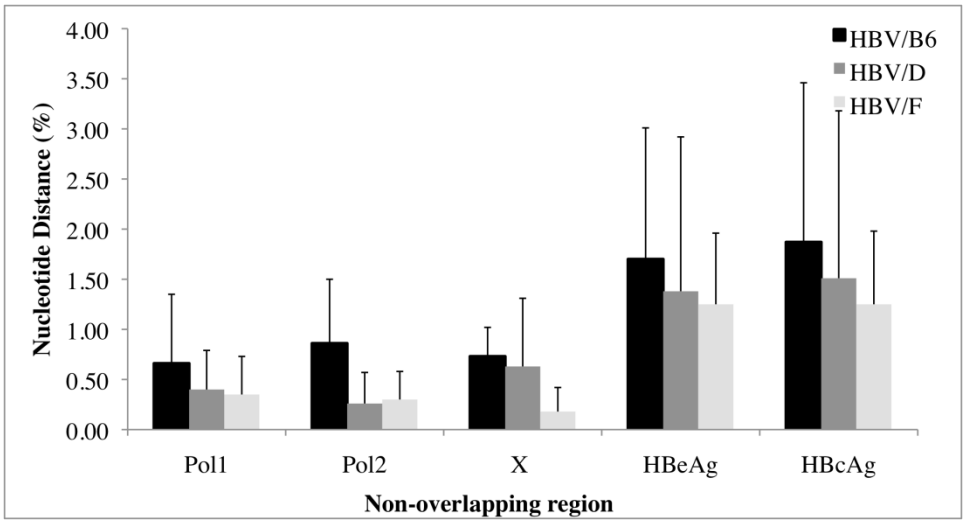


Figure 12. Percent nucleotide distance for each non-overlapping region of the hepatitis B virus (HBV). FindModel was used to determine the most appropriate alignment model, followed by calculation of pairwise distance amongst each region by MEGA4 software [2]. Non-overlapping regions are defined as follows (numbering is based on the theoretical EcoR1 site); Pol1 (polymerase): nt2456-2843; Pol2: nt833-1373; X (HBV X region): nt1621-1813; HBeAg (HBV e antigen): nt1836-2306 and HBcAg (HBV core antigen): nt1901-2306. The HBsAg (HBV surface antigen) was not included as it overlaps completely with the polymerase region. Standard error bars are shown.

Owing to its overlapping nature, the average distance within the HBsAg was found to be three times lower than the non-overlapping region of the HBcAg for all genotypes (0.46%±0.45% vs. 1.54%±1.33%) and was highest within the B6-patients

(0.53%±0.39%, data not shown). Investigating only non-overlapping regions of the HBV genomes further demonstrates that HBV/B6 has the highest distance at the nucleotide level when compared to the other two genotypes (D and F).

The genetic diversity of the non-overlapping regions at the AA level was also investigated and the results are shown in Fig. 13. Although the nucleotide distance is higher within the non-overlapping core region for genotype B6 sequences, the percent nucleotide distance is comparable among the 3 genotypes in this region (ranging from a mean of 1.25%-1.87%). However, the percent amino acid diversity is considerably higher for genotype B6 (8.0%±1.5%) within the non-overlapping core coding region in comparison to genotype D and F amino acid sequences.

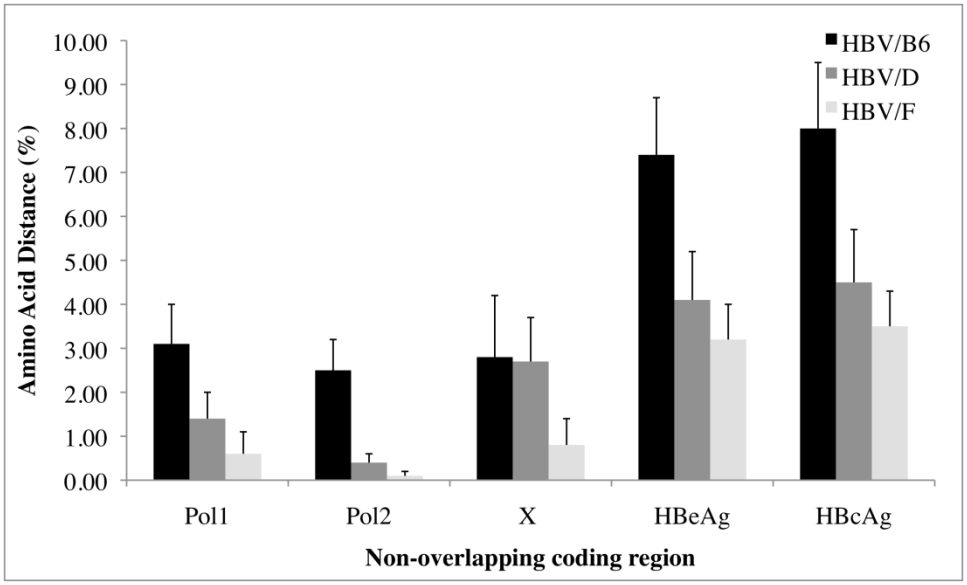


Figure 13. Percent amino acid diversity for each non-overlapping coding region of the hepatitis B virus (HBV). FindModel was used to determine the most appropriate alignment model, followed by calculation of the pairwise distance within each coding area by MEGA4 software[2]. Non-overlapping regions are defined as; Pol1 (polymerase):49-180; Pol2: 582-760; X: 84-159; HBeAg (HBV e antigen): 9-163 and HBcAg (HBV core antigen): 1-135. The HBV surface antigen (HBsAg) was not included as it completely overlaps with the polymerase region. Standard error bars are shown.

The AA diversity of the HBsAg reading frame was investigated and although it was highest amongst the HBV/B6 patients (1.38%±1.38); this region was still less diverse than any other region included in the analysis for the B6 patients. Similar to what is seen at the nucleotide level, HBV/B6 has the highest AA diversity amongst all non-overlapping genomic regions.

The change in AA diversity over time was investigated, owing to the known high error rate of the virus' polymerase, and hence its rapid evolutionary rate (Fig. 13). While the previous figures depict the overall divergence between genotypes being fairly extensive, Fig. 14 shows the variation between the 5-year periods of each genotype is considerably lower.

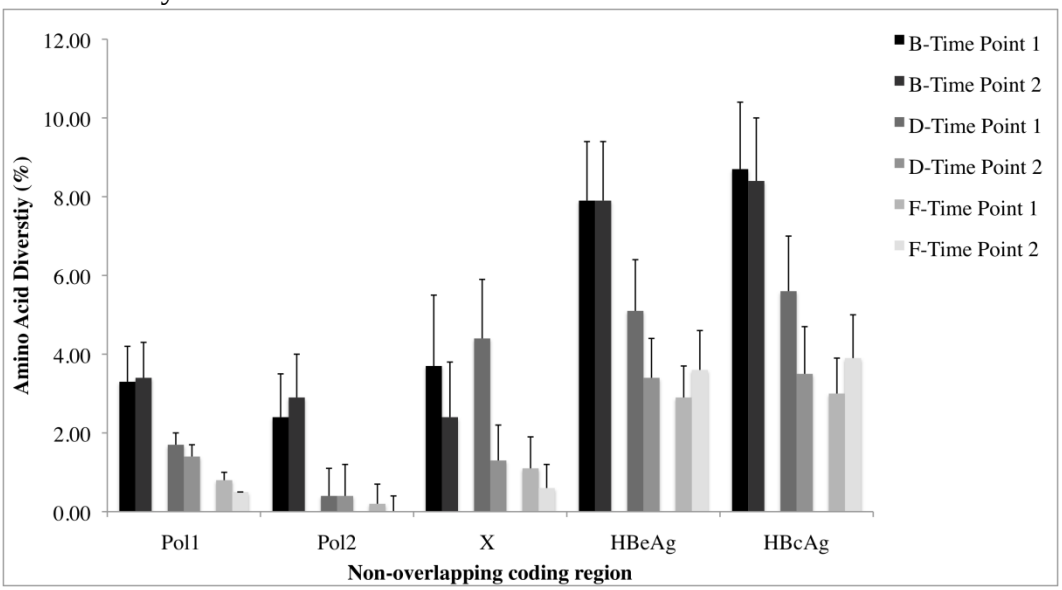


Figure 14. Mean time point amino acid (AA) genetic diversity of each non-overlapping coding region of the hepatitis B virus (HBV). For each genotype, the genetic diversity at the first time point is followed by the second. FindModel was used to determine the most appropriate alignment model, followed by calculation of the pairwise distance within each coding area by MEGA4 software [2]. B: HBV/B6 infected patients, D: HBV/D-infected patients, and F: HBV/F-infected patients. Non-overlapping regions are defined as; Pol1 (polymerase): 49-180; Pol2: 582-760; X (HBV X protein): 84-159; HBeAg (HBV e protein): 9-163 and HBcAg (HBV core protein): 1-135. The HBV surface protein (HBsAg) was not included as it completely overlaps with the polymerase region. Standard error bars are shown.

When incorporating all genetic regions of the virus, including those that are overlapping, the change in AA diversity over time may again be attributed to genotypic characteristics owing to the small change shown (Fig. 15). The PreS and HBsAg regions are included and as expected, these regions exhibit lower degrees of change when compared to other regions and lack any drastic modifications over the study period.

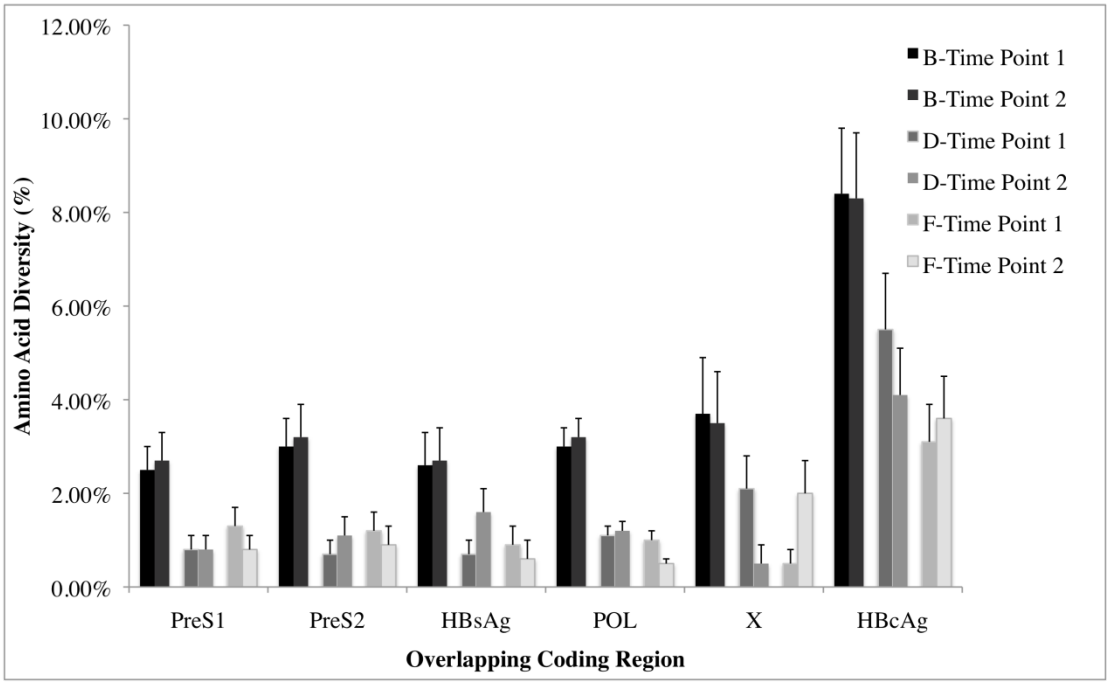


Figure 15. Mean time point amino acid (AA) genetic diversity of each coding region of the hepatitis B virus (HBV). For each genotype, the genetic diversity at the first time point is followed by the second. FindModel was used to determine the most appropriate alignment model, followed by calculation of the pairwise distance within each coding area by MEGA4 software [2]. B: HBV/B6 infected patients, D: HBV/D-infected patients, and F: HBV/F-infected patients. HBsAg: HBV surface protein; POL: polymerase; x: HBV x protein; HBcAg: HBV core protein. Standard error bars are shown.

6. Quasispecies Analysis

The nucleotide and AA diversity within the HBsAg and the non-overlapping HBcAg portion at the quasispecies level was determined using clonal analysis. The following patients were excluded from the HBsAg portion, owing to the inability to obtain a positive cloning reaction: Canada-3a, Alaskan-2a and -5a. To determine the most accurate representation of the quasispecies population within each individual, a minimum of 20 clones were grown, processed and sequenced.

The dynamic evolution of HBV quasispecies over time within the non-overlapping portion of the core region was different amongst the three genotypes (Fig. 16), based on data from one individual, which was representative of all patients within a genotype group. The HBV/B6 genomes maintained a high mutation rate, as depicted in Fig. 16 by the lack of a stable dominant population in either time point. HBV/D shows slight variation over time, but maintains a fairly stable dominant population. Alaskan HBV/F genomes show a lower degree of variation than B6, as indicated by the complete replacement of their dominant quasispecies with another over time. This indicates that although shifts in the F genomic subpopulation occur, the rate of substitution over time is not as constant as B6.

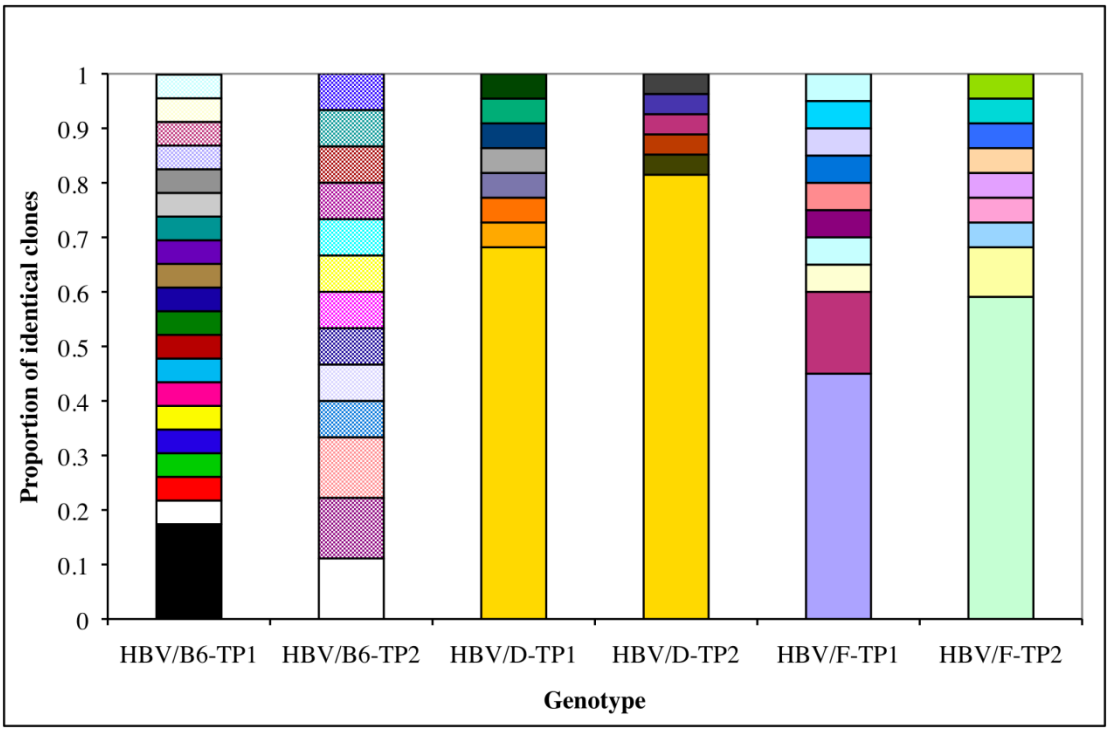


Figure 16. Dynamic evolution of hepatitis B virus (HBV) quasispecies within the non-overlapping region of HBV core antigen (HBcAg) over five years in one representative patient from each genotype. Each colour represents a different quasispecies within each genotype. The same colour in two time points within one genotype indicates identical clonal sequence. For each genotype, the quasispecies population at the first time point (TP1) is followed by the second (TP2). Percent identities were calculated using MegAlign software (v7.1.0; DNAStar Inc., Madison, WI).

The higher proportion of dominant populations in the B6-infected patients indicates lower substitution for the HBsAg region (Fig. 17), compared to the HBcAg region (Fig. 16), meaning this overlapping region evolves slower than the non-overlapping core region and is more functionally constrained. A lower rate of variation within the surface region is further confirmed by the data presented in Fig. 15, whereby the genetic diversity in the surface is lower than the core region for all genotypes at the dominant level.

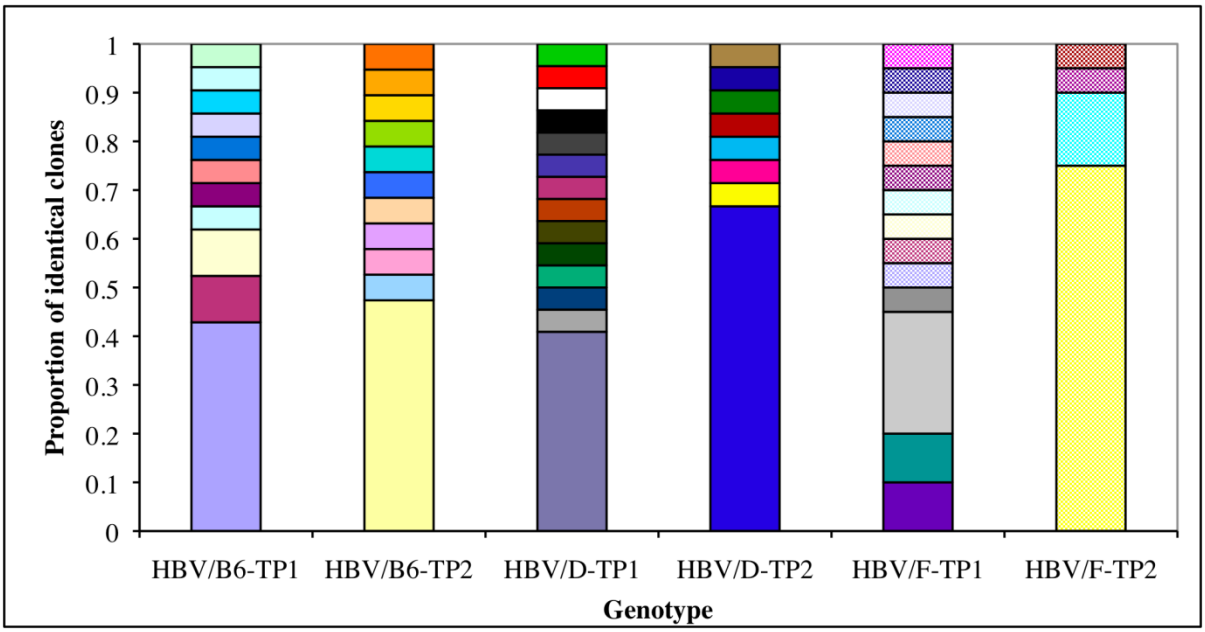


Figure 17. Dynamic evolution of hepatitis B virus (HBV) quasispecies within the HBV surface antigen (HBsAg) over five years in one representative patient from each genotype. Each colour represents different quasispecies within each genotype. The same colour in two time points within one genotype indicates identical clonal sequence. For each genotype, the quasispecies population at the first time point (TP1) is followed by the second (TP2). Percent identities were calculated using MegAlign software (v7.1.0; DNAStar Inc., Madison, WI).

At the quasispecies level, the mean amino acid genetic distance of the non-overlapping core region was much higher amongst all 3 genotypes, when compared to that of the nucleotide level for the core (5.5%±1.1% vs. 2.7%±0.43%) (Fig. 18). The HBV/B6 genomes display a higher genetic distance at both the nucleotide and amino acid levels (4.75%±0.60%; 7.95%±1.4%), when compared to D- (2.40%±0.4%; 4.50%±1.0%) and F genomes (1.05%±0.3%; 3.90%±0.8%). Although not shown, the change over time was negligible for all three genotypes. The amount of quasispecies genetic variation at the amino acid level is higher in all 3 genotypes when compared to that of the nucleotide level.

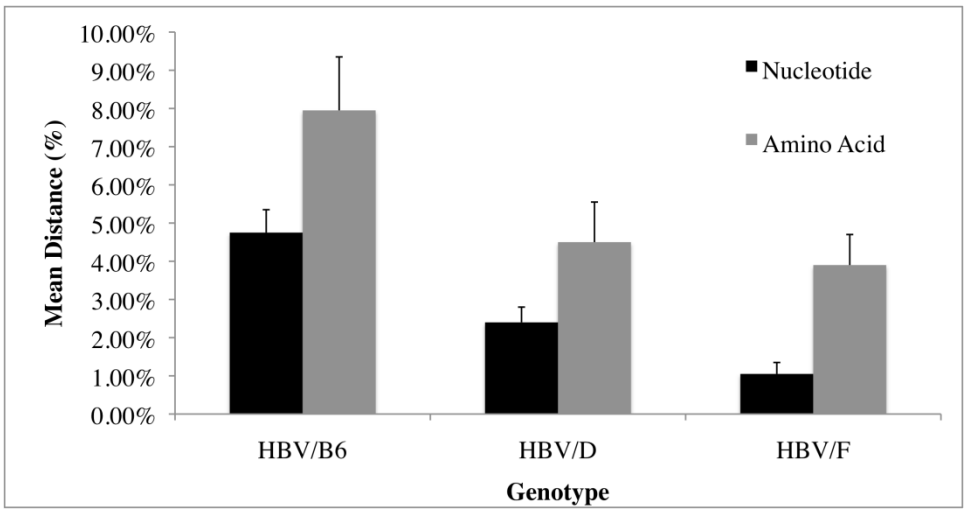


Figure 18. Mean quasispecies genetic distance for all circumpolar hepatitis B virus (HBV) isolates within the non-overlapping core region. The distance within the nucleotide and amino acid coding regions are calculated using Tamura-3 parameter model and p-distance in MEGA4 software [2]. Standard error bars are shown.

The non-overlapping core region at the quasispecies level showed genotypes D and F displayed more nonsynonymous (dN) (3.1%±2.6%; 1.6%±0.4%) than synonymous (dS) changes (2.0%±0.6%; 1.2±0.4%) (Fig. 19). Even though the B6-genomes had more synonymous (dS) changes, the degree of non-synonymous (dN) changes was still drastically higher than both D and F-genomes (4.35%±0.8%). As depicted in Fig. 19, the dN/dS <1 suggests negative or close to neutral selection for B6-isolates, whereas for D- and F-isolates, dN/dS >1 indicates positive selection. In agreement with the previous data, there was no change over time in these values, and the trend remained whereby HBV/B6 strains displayed the largest mean difference in nucleotide changes, when compared to HBV/D- and HBV/F strains.

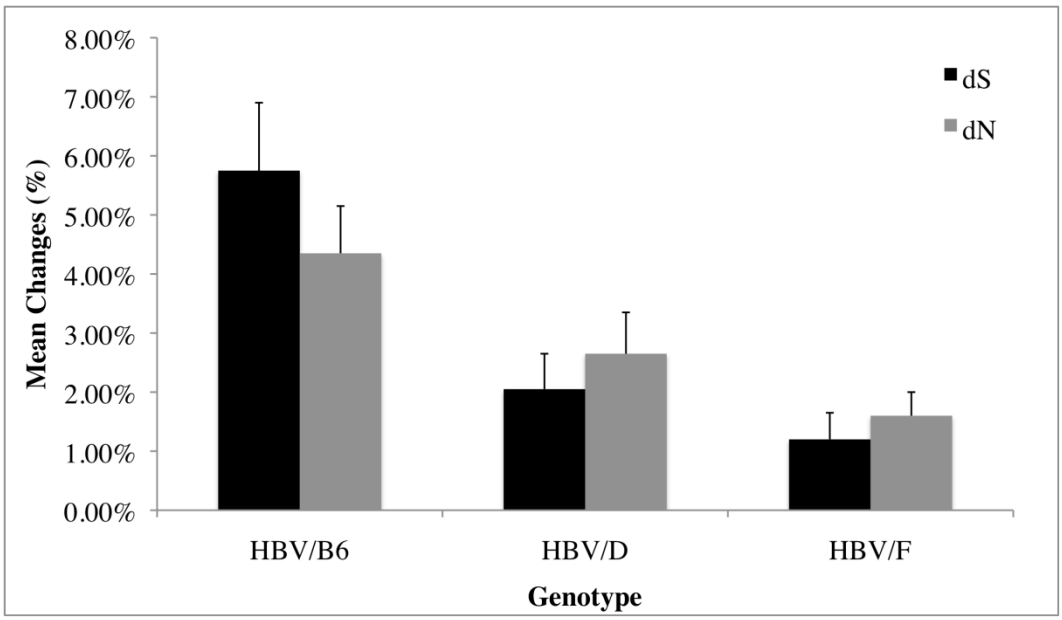


Figure 19. Mean quasispecies synonymous (dS) and nonsynonymous (dN) changes for all circumpolar hepatitis B virus (HBV) isolates within the non-overlapping core region. Changes were calculated using the modified Nei-Gojobori model with Jukes-Cantor correction in MEGA4 software [2]. Standard error bars are shown.

6.1 Correlation between quasispecies nucleotide diversity and viral load

A putative correlation between nucleotide diversity and HBV viral load has been investigated previously, with conflicting results [78, 95]. As depicted in Figs. 20 and 21, this trend was determined using statistical analyses and was found to vary between genotype. Within the non-overlapping HBcAg region (quasispecies level), HBV/B6 and HBV/D specimens are demonstrated in Fig. 20 and show a weak negative correlation between nucleotide diversity and viral load, as depicted by the low R^2 value. Within the B6- and D-quasispecies populations, a decreasing level of HBV DNA was shown to lead to an increase in the nucleotide diversity. A different scenario is depicted for HBV/F. With specimens from this cohort, all of whom eventually developed HCC, we see that with

increasing viral load, the nucleotide distance within the core region also increases as depicted by the weakly positive correlation in Fig. 21.

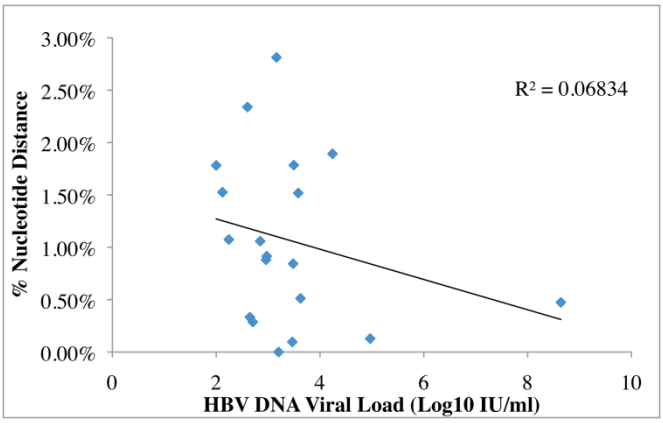


Figure 20. Negative correlation between hepatitis B virus (HBV) DNA level (log_{10} IU/ml) and HBcAg (HBV core antigen) nucleotide distance within the quasispecies population of HBV genotype B6- and D-infected patients (Pearson's correlation; p = .278). The R-squared value is indicated on the graph. Nucleotide distance is calculated using FindModel and MEGA4 software for the pairwise distance [2]. Viral load measurements were taken using Qiagen *artus* HBV TM PCR kits (Qiagen Inc).

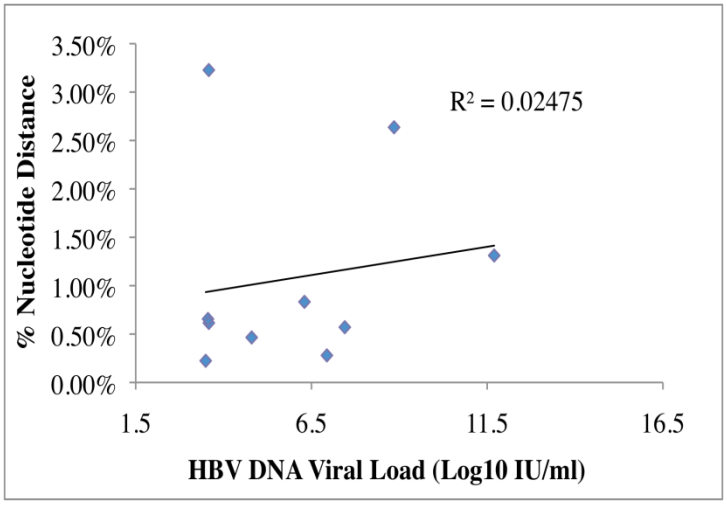


Figure 21. Positive correlation between hepatitis B virus (HBV) DNA level (log_{10} IU/ml) and HBcAg (HBV core antigen) nucleotide distance within the quasispecies population of HBV genotype F-infected patients (Pearson's correlation; p = .667). The R-squared value is indicated on the graph. Nucleotide distance is calculated using FindModel and MEGA4 software for the pairwise distance [2]. Viral load measurements were taken using Qiagen *artus* HBV TM PCR kits (Qiagen Inc).

DISCUSSION

The true extent of HBV infection within the circumpolar regions of the world may be underreported as the majority of studies are based upon a single serum marker (HBsAg). In this study, the prevalence of OBI within a Northern Canadian population was determined by testing for an additional marker of HBV infection, viral DNA. To the best of our knowledge, this work also provides one of the first analyses of longitudinal paired specimens from patients of similar ethnicities [96-97] from remote regions, and infected by one of three HBV genotypes with varying outcomes. This study contributes to an understanding of the potential association between clinical outcome and viral variation in HBV infections.

Based on a previous study by Minuk *et al* (2005) and considering for false-positives, the hypothesized prevalence of OBI was 8-10% [72]. The determined prevalence in our study (3.8%) was therefore lower than expected. This may have been the result of involving more communities and also involving exclusively non-Inuit persons, when compared to this previous study by Minuk *et al* (2005) which included only Inuit persons [72]. The occurrence of OBI is 3.8% in this Northern Canadian community and as this event is more common than HBsAg-positive HBV in Canada (2%) [98]. A high prevalence of OBI is found in HBV endemic regions or in populations with a greater risk of HBV transmission, such as IV drug users and patients co-infected with HIV and HCV [69-70]. The prevalence of HBV infections within Northern Aboriginal communities in Canada is estimated at 3-5% (based on HBsAg positivity), with 14.7-27% having evidence of previous exposure to HBV [79, 99-100]. The stringent definition of what constitutes a positive

OBI used here (HBV DNA sequence confirmed in one region of the virus or PCR positive in two viral regions) may cause the determined prevalence to be underreported as other areas have reported the prevalence of OBI to be 0-17%, also analyzing serum HBV DNA [101-103]. The lower than expected prevalence of OBI may be owing to the difference amongst genotypes of HBsAg-positive and –negative individuals.

The most prevalent genotype here is HBV/A (48.1%), which contrasts with a study done on inhabitants of the Canadian Arctic, where HBV/B and HBV/D prevailed amongst HBsAg-positive individuals [17]. From this previous study, a higher degree of genotypes B and D are expected within the OBI population, yet this discrepancy may be due to the short PreC/C region being the only viral region amplified and sequenced in the majority of the HBV/A cases. For genotyping HBV, whole genome sequencing is widely considered the "gold standard", although using the single HBsAg gene is an adequate and more practical alternative for genotyping [104-105]. It is commonplace for occult-HBV studies to genotype only one region, most often the HBsAg [106-107], as it is known to provide a more accurate genotype determination compared to that of the HBcAg coding region [83]. Genotyping reliability is also dependent upon the size of the sequence analyzed [105]. An example of this is shown in our study, whereby using two different genomic regions of HBV for genotyping can lead to discrepant results, such as with occult patient #13. The occult-HBV isolate from this patient clustered with HBV/B6 (based on the HBsAg sequence); yet with genotype C1 using sequence from the core region (Fig. 7). In this case, the length of the HBsAg sequence used in genotyping was almost double that of the HBcAg sequence, thus allowing for a more accurate phylogenetic comparison (313 nt for HBsAg vs. 177 nt for HBcAg). An Indian study amplified the core and surface

regions to determine the OBI prevalence, yet only used the surface region for genotyping, as perhaps they found it more accurate [108]. Re-amplification of the OBI-positive HBV/A isolates is currently being performed, including amplification of surface gene regions, and may reveal an alternative genotypic distribution once sequence is available.

Identification of potential risk factors for OBI would be highly useful, and has yet to be studied to any great extent in the non-blood donor population of Canada. As expected, almost all OBI-positive individuals were of First Nations ethnicity since the population of Aboriginal peoples in this region is approximately 90% [109]. Similarly, a high distribution of Inuit peoples were seen in the study by Minuk *et al* (2005) [72]. Interestingly, almost all of the occult-positive patients resided within the same settlement (81.5%) and may indicate a possible transmission source, although the phylogenetic data presented here would suggest otherwise from the lack of tight clustering. The gender, age and anti-HBs status did not assist in distinguishing between occult HBV-positive and –negative patients. This was also seen in a study investigating the prevalence of OBI in HCC patients [85]. Although twice as many females were OBI-positive as males, the significance between occult infections and gender has yet to be confirmed. Some research shows a higher proportion of males are OBI-positive, however most of these studies have limited sample size [68]. The presence of Anti-HBs indicates either previous HBV immunization or recovery from an HBV infection, and this marker is less common in OBI-positive patients [68-69, 72]. Within all anti-HBs-positive patients in the current study, only 4% (11/274) were also HBV DNA positive and is in agreement with previous research [72].

The presented OBI prevalence may be more valid than previous studies owing to the use of nested PCR. This method of PCR is believed to be more sensitive than single stage PCR for detecting OBI, due to the low level of template DNA [69, 110]. Selecting sensitive primers and sequencing the resulting amplicon increase specificity in detecting OBI patients. The primer sets used in amplification of the PreC/C (HBV5/6 and HBV7n/8n) region have a previously reported lower limit of detection between 10^3 and 10 genome equivalents [85] and 0.1-0.05 IU/ml [111]. Tests of sensitivity completed in-house were in agreement with the Pollicino study from 2004 (unpublished data) [75]. Whether amplification of either the HBsAg or PreC/C regions is more efficient is dependent on the particular primer set used, yet the greater the sequence length of the amplicon, regardless of the region amplified, the greater the accuracy of genotyping. The method of choice for DNA extraction in the detection of OBI also has yet to be optimized, although amplification of low-level viral DNA has been found to be most efficient using the method employed here (phenol chloroform) [80].

The "gold standard" for determining occult infection is the detection of HBV DNA in the liver of HBsAg-negative individuals, yet it is obviously difficult to obtain liver biopsy samples in the case of an asymptomatic patient [110]. As such, analysis of blood extracts is considered a sufficient alternative and is employed here [110]. Although the virus normally replicates within hepatocytes, blood is used as a non-invasive means of diagnosis, even in HBsAg-positive virus detection. The presence of HBV DNA in serum is an indirect indication of replication, yet studies show occult-HBV cases do possess low-level replication within peripheral blood mononuclear cells and other extrahepatic locations [69].

The clinical importance of OBI is debated [112-113], yet many studies link these infections to adverse outcomes, such as a risk factor in the development of HCC [85], causing activation of seemingly asymptomatic HBV infection [114], and clinical aggravation of HCV-infected patients [112]. These studies provide evidence that OBI infections should not be ignored, especially in an ethnicity with a higher than average risk for HBsAg-positive HBV infections. This studied population is an example of an endemic area displaying OBI, with a potential for adverse outcomes if unmonitored.

Occult HBV infections represent a field of study being increasingly researched, aided by more sensitive detection techniques. Focus remains on the clinical management of HBV and this study attempts to further define the association between clinical outcome and HBV genetic variation.

Previous studies comparing those with adverse vs. benign outcomes have been difficult owing to the lack of paired, longitudinal samples to investigate an additional area of confusion: the change over time of HBV infection. In addition, studies investigating one geographical area pose the risk of enrolling patients with different immunological profiles, if the patients reside in a multi-ethnic city. In an attempt to control for host effects, such as immunological profiles on HBV infections, this study has attained samples from individuals of only Inuit/Alaska Native ethnicity. Archaeological research suggests ancestors of the Thule peoples crossed the Bering land bridge more than 10,000 years ago and included ancestors of modern Alaska Natives, and Canadian and Greenlandic Inuit [97]. This population is believed to have expanded further to the east, into Canada and eventually Greenland [96-97]. This evidence, along with mitochondrial DNA

studies suggests the western circumpolar Aboriginal groups share a related ancestry and a common migratory event, more than 10,000 years ago [115].

In the genetic characterization of HBV in circumpolar indigenous populations, patients infected with HBV/B6 have yet to be associated with any adverse clinical outcomes [4, 16]. The patients infected with HBV/B6 were all noted to have levels of HBV DNA <4.39 $\log_{10}$ IU/ml, which has not been associated with a high risk of developing HCC [116]. This is in contrast to those infected with HBV/F, as all patients in this study are eventually diagnosed with HCC and have viral loads >4.39 $\log_{10}$ IU/ml [116]. Cancer of the liver has also been associated with patients infected with HBV/F previously [14]. The HBV/D-infected Greenlandic Inuit are assumed to have no liver-related adversities during the period of study (M. Borresen, personal communication). Similar to the B6 isolates, the viral load of HBV/D is also indicative of a lower risk of liver disease, as four of five HBV/D-infected patients had levels <4.39 $\log_{10}$ IU/ml [116]. Older age is also frequently associated with many factors in HBV infections, including HBe-SC [78], the development of cirrhosis and HCC [117] and some mutations [118]. As patients infected with HBV were not age-matched in this study, this presents a limiting factor. Owing to the young average age of those infected with HBV/D in this study (33.2 years), one may not ignore the future potential of adverse outcomes related to HBV infections.

The presence of specific mutations within the HBV genome has been previously associated with adverse outcomes [18-20, 46, 48, 54-55, 57]. The results of this study show while these substitutions are found in all implicated genotypes; HBV/F is the only to have deletions. These deletions are located within the PreS region, where they have been associated with HCC [118] and are found in the two HBV/F pairs also displaying the highest

level of HBV DNA. In particular, this region is known for its hypervariability, as viral particles completely lacking a PreS are still infective [119]. These same two HBV/F pairs also display deletions within the HBx protein, another common finding in HBV-infected HCC patients [120]. As expected, the common BCP mutant, A1762T/G1764A (TA) [121], was found in all HBV/F pairs, and in none of the HBV/B6 pairs. Reports show the TA mutant has a significant association with HCC; even after the adjustment for genotype and viral load [122]. Although a large difference exists between the average age of the B6- and F-infected patients (66 vs. 30.6 years at baseline), the study by Liu *et al* (2009) demonstrates after correcting for age and duration of chronic HBV infection, mutations within the core promoter are still associated with HCC [123]. These mutants are selected earlier in HCC development, in comparison to other core promoter mutations (T1766/A1768), which arise later according to longitudinal data [124] and as shown by this study.

A mutation associated with the inactive carrier state of chronic HBV infections, G1896A, abolishes production of the HBeAg [125], and its appearance is normally associated with HBe-SC[126]. Known for a bias towards genotypes B, D and E [126], this mutant is found in all HBV/B6 genomes and most HBV/D pairs here. HBeAg is a common immunoregulatory protein [126], and its absence is expected in benign HBV cases, as this helps to prevent immune recognition, and hence immune-related damage. The association between the 1896A mutant and inactive carriers has been previously found in HBV/B-infected circumpolar indigenous populations [10, 17, 127-128]. A discrepancy exists for HBV/D pair 2, as it possesses the mutant in the first time point, yet is also HBeAg positive. Residual HBeAg is occasionally observed in patients infected with G1896A

variants due to the non-specific detection of partially degraded nucleocapsids released from infected cells [126]. The presence of the G1896A mutant in this study is associated with benign, inactive HBV carriers and is demonstrated in this study with the detection in HBV/B6-isolates

This study shows the benign outcome-associated HBV/B6 genome has a higher overall diversity than HBV/D and HBV/F (4.1% vs. 1.3% and 0.7%, respectively). Over the full-length nucleotide sequence, the B6 genome is significantly more diverse, despite a lower average viral load when compared to the other two genotypes. All patients infected with HBV/B6 were HBeAg negative and are assumed to have HBe-SC decades earlier. It has been shown that HBeAg seroconverters become more genetically varied following conversion compared to non-seroconverters [129]. During HBeAg seroconversion, the host's adaptive immune system responds to viral antigens, such as HBeAg, driving HBV mutation rates [95] and lowering the HBV DNA level, as depicted by the decrease in viral load for B6 isolates, yet increase in nt diversity (Fig. 20). HBV associated with HBeAg negative infection evolves more rapidly and has more overall variation than those in HBeAg positive infection, such as the case of B6 vs. D- and F-isolates [130-131]. In addition, research shows the evolutionary rate of HBV is higher in HBeAg negative than positive patients [16].

Those infected with genotype D or F comprised both HBeAg-negative and –positive patients. In contrast, all B6-patients were HBeAg-negative. The lack of biochemical and clinical evidence of active liver disease in the B6-infected patients may indicate a weaker immune response. CD8+ T-cells specific to the HBcAg predominate within the total CD8+ T-cell population during HBV infection, and are important for

controlling infection, yet also contribute to the development of liver disease [132]. Thus, in chronically B6-infected patients who lack evidence of liver disease, the HBV-specific T-cell response is likely weak, reflecting T-cell tolerance and allowing for viral persistence [133], thereby removing a selective force. Within the non-overlapping HBcAg region, the B6-isolates fell under slight negative and/or neutral selection, as demonstrated by a dN/dS<1 (Fig. 19). Negative and/or neutral evolution predominates during HBV evolution [16, 134], in part owing to the extensive overlapping of open reading frames [135]; yet negative selection mediated by the virus functions during HBV evolution, depending on the phase of disease [95].

Non HBe-SC, or recent converters are also noted to display a relatively high viral load and low quasispecies diversity, in comparison to converters [129]. Two HBV/D patients did not seroconvert (pair 1 & 2) and are found to have the highest viral loads of all D-infected pairs, whereas, the three patients whom had previously HBe-SC had, on average, lower viral load similar to that of B6-infected patients. HBV/F-infected patients had the highest level of HBV DNA and the lowest quasispecies diversity within the HBcAg and HBsAg regions (Fig. 16 and 17), further confirming the findings of Wu *et al* (2011) [129]. This is in contrast to HBV/B6 since B6-isolates had higher HBcAg quasispecies variation (4.75%±0.6% vs. 2.40%±0.3% [HBV/D] and 1.05%±0.3% [HBV/F]) and lower average viral load than both D and F genomes (2.96 vs. 3.83 and 5.01 $\log_{10}$ IU/ml, respectively). Higher rates of genetic change have been previously related to HBV virion generation time [95]. The idea of a shorter generation time in carriers leading to genetic variation is also proposed in a mathematical model by Nakabayashi & Sasaki (2011) [136]. They propose the “arrested” or inactive virus evolves in the event of

rapid virion release. Subgenotype B6 may be capable of quicker generation time than D and F, leading to higher degrees of genetic variation while maintaining a low viral load. Thus, HBV/B6-isolates demonstrate higher levels of genetic variation than D- and F-isolates, possibly due to the phase of the disease and a reduced generation time.

The elevation in nucleotide diversity is persistent over time with B6, yet the B6-infected patients lack biochemical and clinical evidence of active infection, involving immune activation [10, 16, 127]. HBV/B6 isolates show evidence of negative selection within the non-overlapping core coding region, ($dN/dS<1$; Fig.19), when compared to D and F-isolates which demonstrate a slight positive trend in this region (Fig. 19). The average age of the B6-infected patients may have allowed for a longer duration of infection, during which time substitutions accumulated and intrapatient viral evolution increased. The B6 viruses have therefore undergone a period of increased nucleotide substitution to expand its QS population (Fig. 16), which D- and F-isolates lack. Considering the age and the phase of disease of most D- and F-infected patients, nucleotide changes resulting in nonsynonymous amino acid changes within the HBV/D and F viruses is consistent with immune selective pressures resulting in mutations allowing for escape of immune recognition [95]. The younger average age of these patients indicates they may have been infected for a shorter period than B6, and hence less viral evolution.

Research shows the persistent accumulation of HIV within isolated communities leads to a unique adaptation from population-specific selection pressures [137]. Infections with subgenotype B6 are found in individuals of Inuit descent inhabiting the western circumpolar regions [127], many of which live their entire lives in remote, isolated communities [138]. Thus, adaptation of HBV/B6 to this particular host may play a role in

the type of infection and disease outcome experienced by these northern Canadian Inuit. The increased genetic distance and viral quasispecies diversity observed with HBV/B6 strains may originate in the diversifying selection driven during loss of immunotolerance or have been inherited from the donor. Thus, these populations are subject to unique selection pressures, which have shaped a possible host-pathogen balance. The balance arises from a degree of immune control, demonstrated by the low level of HBV DNA, yet the B6-virus escapes immune recognition to allow persistence with a lack of immune-mediated liver damage.

A theory has been proposed for the possible source of HBV/B6 in the circumpolar region [127]. The study suggests that the Aboriginal people migrating from Siberia to Alaska across the Bering land bridge 10,000 years ago may have been infected with a common Asian variant of HBV, subgenotype B1. The population migrated further east towards Greenland in the last 1000 years making it possible subgenotype B6 in the Inuit population has evolved from B1. The long history this subgenotype may have with the population may make for a less virulent course of disease, as depicted by the asymptomatic Canadian Inuit here. Pathogen attenuation is a term used to describe the coevolution of host and pathogen, resulting in virulence reduction [139]. The occurrence of such an event is believed to have taken place in myxoma virus infection in rabbits [140] and within human cases of syphilis [141] and possibly HIV-1 subtype C [139].

Genotypes HBV/D and HBV/F may have been more recently introduced to the circumpolar indigenous peoples, when compared to HBV/B6 and thus, these two may have less evolutionary history, hence less adaptation allowing for more harmful outcomes. Evidence for this recent event is supported by the high rate of HCC in the

Alaska Natives [14], potentially caused by a lack of virulence attenuation. Based on phylogenetic analyses, previous work shows HBV/ F1 from Alaskan HCC patients clusters with that of Argentinean strains of HBV/F1 and may suggest a common ancestor [142]. Phylogenetic analysis has also been used previously by Central American researchers to speculate on the evolution of HBV/F, owing to the high prevalence rate of genotype F infections in their respective countries [143]. The limited phylogenetic information available is compatible with the hypothesis that HBV/F was introduced into Alaska Natives more recently than HBV/B6 into the Canadian Inuit. Preliminary data for the most recent common ancestor (MRCA) of HBV genotype D2 is approximately less than 100 years, although more accurate estimations require a larger number of samples [142]. The history of genotype D within the Inuit of Greenland may have coincided with the emergence of HBV/F in Alaskan Natives, approximately 500-1,000 years ago. The lack of clinical information on the HBV/D-infected Inuit in this study limits any evolutionary conclusions from being made on this genotypic group. Although conclusions for the evolutionary history of genotypes B6, D and F-infections within these circumpolar groups cannot be made, it is clear that HBV/B6 may have a long-term history in Inuit of the Canadian arctic, and thus a possibly less virulent course of disease when contrasted with D- and F-genomes, which may be more recently introduced to their respective populations.

Dual functioning reading frames place constraints on their genomic evolution; hence studying non-overlapping regions may assist in investigating the true effect of selection pressure. Here, the non-overlapping portion of the PreC/C region displays the highest diversity at the nt and AA level and amongst all genomic regions. It has been

suggested HBeAg negative carriers have a higher degree of AA variation within the HBcAg, when compared to HBeAg-positive carriers [131]. This is confirmed, as the genotype with the most diverse PreC/C region was that of subgenotype B6 (1.87%±1.59%). The presence of the PreC stop codon (G1896A) was also shown to be highly associated with the frequency of core mutations [144-145] in HBV chronic carriers, also shown by B6 genomes in the current study. PreC/C mutants arising from immune selection is suggested to occur around the time of seroconversion, as shown by Milich & Liang (2003) [126]. This may generate high variation within the PreC/C region, during the HBeAg-negative state. This is also seen in the case of HIV and HCV, where specific genetic regions are shown to vary more than other regions, particularly in common immunological epitopes [146-147]. Investigating the non-overlapping core region allows us to partially control for evolutionary restraints posed on the overlying genomic regions of HBV. Here, we show the non-overlapping core region is highly variable when compared to all other genomic regions, especially in the case of HBV/B6 genomes.

Constrained evolution within the HBsAg region has been previously suggested, owing to the completely overlapping nature with the POL ORF [16, 135]. At the dominant and subdominant level of all genotypes involved in this study, the HBsAg region shows lower variation than the non-overlapping HBcAg region (Fig. 15 & 17) (0.46%±0.45% vs. 1.54%±1.33% at the dominant level). This is in agreement with previous understanding, as mutations here are kept to a minimum owing to the overlap with the POL region, making any variations potentially fatal for the virus [135].

The characteristic life cycle of HBV is marked by a high error rate due to its DNA polymerase, as it lacks proofreading capabilities. The viruses' evolutionary rate has been

studied, and from a 25-year longitudinal study on inactive carriers, the mean number of nucleotide changes per year was calculated to be slightly higher than previous estimates at approximately $7.9x10^{-5}$ [16]. For this reason, the change in genetic variation over time is included in this current study, as one may expect relatively high degrees of change, when compared to other DNA viruses. At both the dominant and quasispecies levels, the B6 genomes are found to maintain a high mutation rate over the five-year period (Fig. 14-17). Within the subpopulations, the complete replacement of the dominant quasispecies over time in HBV/B6 also indicates this subgenotype has an inherently high rate of substitution when compared to that of HBV/D and HBV/F. For those infected with HBV/F or HBV/D, the viral genomes experience shifts in the genomic dominant and quasispecies populations, yet the dynamic evolution was not as pronounced as that observed with B6. Based on the phylogenetic tree (Fig. 8), HBV/F and HBV/D isolates have less inherent difference between pairs as shown by the shorter branch lengths, in comparison to HBV/B6 pairs. This suggests, along with the insignificant amount of diversity within each genotype (Fig. 11), that the high degree of genetic variation of B6 is characteristic of this subgenotype. The period studied was five years and this was not considered a sufficient period of time to accurately determine a molecular evolutionary rate.

Conclusions

In summary, the understated scope of HBV infections within Northern Canadian populations owes in part to a lack of study and possible insufficiencies in detection. In rejecting the hypothesized occurrence, this study reports the prevalence of OBI in the studied population to be higher than the national average prevalence of HBV (HBsAg-positive) infections. It therefore stands to reason that further investigation into the clinical significance of OBI is warranted, as current research implies they are potentially harmful.

The inherently high genetic variation within an inactive carrier of HBV is related to a multitude of factors, including shorter virion generation time. All B6-isolates were HBeAg-negative suggesting the high variation is related to immune activation following the loss of immunotolerance. The results suggest rejection of the second hypothesis whereby higher variation would be found in genotypes associated with adverse outcomes. The B6-genomes may be subject to unique forces, such as diversifying or negative selection that likely created a host-pathogen balance. In contrast, HBV/F-genomes were unable to attain this balance between host and pathogen, as indicated by the incidence of HCC in the F-infected patients resulting from immune-mediated damage.

In addition, this study shows deletions found in HCC-associated genotype F are more associated with adversity, than that of specific substitutions. Findings related to the association between clinical course and HBV/D infection are less clear owing to the younger average age of these patients and the lack of clinical data. It would stand to reason based on the results presented, that the characteristics of HBV/D infections lie between the spectra of HBV/B6 and HBV/F. The history of HBV/B6 in circumpolar Inuit

is not well understood, although phylogenetic analysis has provided some evidence. One cannot ignore the possibility of viral adaption owing to the benign clinical course of B6 infections vs. the adverse outcome with HBV/F.

While the use of HBV genotypes in predicting patient outcome is still far from clinical management utility, the developments here can presumably lend themselves to many future applications. Areas that serve to benefit from these findings include personalized treatment choices and the development of preventative therapies. Patients recently diagnosed with HBV may undergo testing for the infecting genotype in hopes of providing a more targeted course of therapy. Preventative therapies may also benefit from this study by pinpointing other potential targets for vaccine development.

Future Directions

The detection assay used here allowed for the determination of OBI-positive individuals; however the possibility of false-positives should not be ignored. Therefore, obtaining follow-up samples from these patients would confirm their positive result. In addition, follow-up samples would provide longitudinal information on these infections to further define the clinical outcome associated with OBI.

The relationship between clinical outcome and genetic variation in HBV is of continued interest and while the results here assist in defining this association, follow-up samples and more extensive clinical data would provide further evidence. Also, this would eventually provide sufficient data for the calculation of the molecular evolutionary rate of HBV. Studies involving larger sample sizes are also needed. In order to study the effect of positive selection within HBV, one would need to obtain the human leukocyte

antigen (HLA)-type of all individuals, such that conclusions can be made on specific epitopes and/or areas that are under selective pressures. Having all study subjects with the same HLA type would allow enhanced control of host factors in the disease course of HBV to define the virus's role. Phenotypic assays are lacking in this area of study, and in order to further define the role of genetic variation has in clinical outcomes, *in vivo* studies must be performed.

References

1. Deng W, Maust BS, Nickle DC, et al. DIVEIN: a web server to analyze phylogenies, sequence divergence, diversity, and informative sites. *BioTechniques.* 2010;48(5):405-408.

2. Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular biology and evolution.* 2007;24(8):1596-1599.

3. Rehermann B, Nascimbeni M. Immunology of hepatitis B virus and hepatitis C virus infection. *Nat Rev Immunol.* Mar 2005;5(3):215-229.

4. McMahon BJ. Natural history of chronic hepatitis B. *Clinics in liver disease.* 2010;14(3):381-396.

5. Liaw YF, Chu CM. Hepatitis B virus infection. *Lancet.* 2009;373(9663):582-592.

6. Kurbanov F, Tanaka Y, Mizokami M. Geographical and genetic diversity of the human hepatitis B virus. *Hepatology research : the official journal of the Japan Society of Hepatology.* 2010;40(1):14-30.

7. Seeger C, Zoulim F, Mason WS. Hepadnaviruses. In: Knipe DM, Howley PM, eds. *Field's Virology.* Vol 2: Lippincott Williams & Wilkins; 2007:2977.

8. Arctic Council. Inuit Circumpolar Council (ICC). 2011; http://arctic-council.org/participant/icc. Accessed March 10, 2011.

9. Tulisov A, McMahon BJ, Koch A, et al. Viral hepatitis in the Arctic. A review from a Circumpolar Workshop on Viral hepatitis, ICCH13. *Alaska Medicine.* 2007;49(2 Suppl):193-203.

10. Krarup HB, Andersen S, Madsen PH, Okkels H, Hvingel BH, Laurberg P. Benign course of long-standing hepatitis B virus infection among Greenland Inuit? *Scandinavian journal of gastroenterology.* 2008;43(3):334-343.

11. Mackie CO, Buxton JA, Tadwalkar S, Patrick DM. Hepatitis B immunization strategies: timing is everything. *CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne.* 2009;180(2):196-202.

12. Harpaz R, McMahon BJ, Margolis HS, et al. Elimination of new chronic hepatitis B virus infections: results of the Alaska immunization program. *J Infect Dis.* Feb 2000;181(2):413-418.

13. Borresen ML, Olsen OR, Ladefoged K, et al. Hepatitis D outbreak among children in a hepatitis B hyper-endemic settlement in Greenland. *Journal of viral hepatitis.* 2010;17(3):162-170.

14. Livingston SE, Simonetti JP, McMahon BJ, et al. Hepatitis B virus genotypes in Alaska Native people with hepatocellular carcinoma: preponderance of genotype F. *The Journal of infectious diseases.* 2007;195(1):5-11.

15. Hurlburt KJ, McMahon BJ, Simonetti JP, et al. Hepatitis B-associated vasculitis in Alaska Natives: viral genotype, clinical and serologic outcome. *Liver international : official journal of the International Association for the Study of the Liver.* 2007;27(5):627-632.

16. Osiowy C, Giles E, Tanaka Y, Mizokami M, Minuk GY. Molecular evolution of hepatitis B virus over 25 years. *Journal of virology.* 2006;80(21):10307-10314.

17. Osiowy C, Larke B, Giles E. Distinct geographical and demographic distribution of hepatitis B virus genotypes in the Canadian Arctic as revealed through an extensive molecular epidemiological survey. *Journal of viral hepatitis.* 2010(Journal Article).

18. Chotiyaputta W, Lok AS. Hepatitis B virus variants. *Nature reviews.Gastroenterology & hepatology.* 2009;6(8):453-462.

19. Song BC, Cui XJ, Kim HU, Cho YK. Sequential accumulation of the basal core promoter and the precore mutations in the progression of hepatitis B virus-related chronic liver disease. *Intervirology.* 2006;49(5):266-273.

20. Sung JJ, Tsui SK, Tse CH, et al. Genotype-specific genomic markers associated with primary hepatomas, based on complete genomic sequencing of hepatitis B virus. *Journal of virology.* 2008;82(7):3604-3611.

21. Shepard CW, Simard EP, Finelli L, Fiore AE, Bell BP. Hepatitis B virus infection: epidemiology and vaccination. *Epidemiologic reviews.* 2006;28(Journal Article):112-125.

22. Bloodborne Pathogens Section BSS, Health Care Acquired Infections Division HC. Hepatitis B Fact Sheet. Vol 2009.

23. Dawar M, Patrick DM, Bigham M, Cook D, Krajden M, Ng H. Impact of universal preadolescent vaccination against hepatitis B on antenatal seroprevalence of hepatitis B markers in British Columbia women. *CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne.* 2003;168(6):703-704.

24. Krajden M, McNabb G, Petric M. The laboratory diagnosis of hepatitis B virus. *Can J Infect Dis Med Microbiol.* Mar 2005;16(2):65-72.

25. Hayashi K, Katano Y, Takeda Y, et al. Association of hepatitis B virus subgenotypes and basal core promoter/precore region variants with the clinical features of patients with acute hepatitis. *Journal of gastroenterology.* 2008;43(7):558-564.

26. Lau GK. HBeAg-positive chronic hepatitis B: why do I treat my patients with pegylated interferon. *Liver international : official journal of the International Association for the Study of the Liver.* 2009;29 Suppl 1(Journal Article):125-129.

27. Livingston SE, Simonetti JP, Bulkow LR, et al. Clearance of hepatitis B e antigen in patients with chronic hepatitis B and genotypes A, B, C, D, and F. *Gastroenterology.* 2007;133(5):1452-1457.

28. McMahon BJ. The influence of hepatitis B virus genotype and subgenotype on the natural history of chronic hepatitis B. *Hepatology international.* 2009;3(2):334-342.

29. Brechot C, Kremsdorf D, Soussan P, et al. Hepatitis B virus (HBV)-related hepatocellular carcinoma (HCC): molecular mechanisms and novel paradigms. *Pathol Biol (Paris).* Aug 2010;58(4):278-287.

30. Tennant BC, Toshkov IA, Peek SF, et al. Hepatocellular carcinoma in the woodchuck model of hepatitis B virus infection. *Gastroenterology.* Nov 2004;127(5 Suppl 1):S283-293.

31. Husa P, Linhartova A, Nemecek V, Husova L. Hepatitis D. *Acta virologica.* 2005;49(4):219-225.

32. Benhamou Y. HIV-1/hepatitis B coinfection. *Expert review of anti-infective therapy.* 2005;3(2):229-239.

33. Nguyen DH, Ludgate L, Hu J. Hepatitis B virus-cell interactions and pathogenesis. *J Cell Physiol.* Aug 2008;216(2):289-294.

34. Chisari FV, Isogawa M, Wieland SF. Pathogenesis of hepatitis B virus infection. *Pathol Biol (Paris).* Aug 2010;58(4):258-266.

35. Beck J, Nassal M. Hepatitis B virus replication. *World J Gastroenterol.* Jan 7 2007;13(1):48-64.

36. Chai N, Chang HE, Nicolas E, Han Z, Jarnik M, Taylor J. Properties of subviral particles of hepatitis B virus. *J Virol.* Aug 2008;82(16):7812-7817.

37. Frelin L, Wahlstrom T, Tucker AE, et al. A mechanism to explain the selection of the hepatitis e antigen-negative mutant during chronic hepatitis B virus infection. *J Virol.* Feb 2009;83(3):1379-1392.

38. Cui XJ, Cho YK, Song HJ, Choi EK, Kim HU, Song BC. Molecular characteristics and functional analysis of full-length hepatitis B virus quasispecies from a patient with chronic hepatitis B virus infection. *Virus research.* 2010;150(1-2):43-48.

39. Kramvis A, Kew M, Francois G. Hepatitis B virus genotypes. *Vaccine.* 2005;23(19):2409-2423.

40. Torres C, Leone FG, Pezzano SC, Mbayed VA, Campos RH. New perspectives on the evolutionary history of hepatitis B virus genotype F. *Mol Phylogenet Evol.* Feb 4 2011.

41. Andernach IE, Hubschen JM, Muller CP. Hepatitis B virus: the genotype E puzzle. *Reviews in medical virology.* 2009;19(4):231-240.

42. Osiowy C, Gordon D, Borlang J, Giles E, Villeneuve JP. Hepatitis B virus genotype G epidemiology and co-infection with genotype A in Canada. *The Journal of general virology.* 2008;89(Pt 12):3009-3015.

43. Yim HJ, Hussain M, Liu Y, Wong SN, Fung SK, Lok AS. Evolution of multi-drug resistant hepatitis B virus during sequential therapy. *Hepatology (Baltimore, Md.).* 2006;44(3):703-712.

44. Chu CJ, Hussain M, Lok ASF. Hepatitis B virus genotype B is associated with earlier HBeAg seroconversion compared with hepatitis B virus genotype C. *Gastroenterology.* 2002;122(7):1756-1762.

45. Yin J, Zhang H, Li C, et al. Role of hepatitis B virus genotype mixture, subgenotypes C2 and B2 on hepatocellular carcinoma: compared with chronic hepatitis B and asymptomatic carrier state in the same area. *Carcinogenesis.* Sep 2008;29(9):1685-1691.

46. Yin J, Xie J, Liu S, et al. Association Between the Various Mutations in Viral Core Promoter Region to Different Stages of Hepatitis B, Ranging of Asymptomatic Carrier State to Hepatocellular Carcinoma. *The American Journal of Gastroenterology.* 2010(Journal Article).

47. Ozasa A, Tanaka Y, Orito E, et al. Influence of genotypes and precore mutations on fulminant or chronic outcome of acute hepatitis B virus infection. *Hepatology.* Aug 2006;44(2):326-334.

48. Lin CL, Kao JH. Clinical implications of hepatitis B virus variants. *Journal of the Formosan Medical Association = Taiwan yi zhi.* 2010;109(5):321-325.

49. Ren X, Xu Z, Liu Y, et al. Hepatitis B virus genotype and basal core promoter/precore mutations are associated with hepatitis B-related acute-on-chronic liver failure without pre-existing liver cirrhosis. *Journal of viral hepatitis.* 2010;17(12):887-895.

50. Sanchez-Tapias JM, Costa J, Mas A, Bruguera M, Rodes J. Influence of hepatitis B virus genotype on the long-term outcome of chronic hepatitis B in western patients. *Gastroenterology.* 2002;123(6):1848-1856.

51. Baumert TF, Thimme R, von Weizsacker F. Pathogenesis of hepatitis B virus infection. *World journal of gastroenterology : WJG.* 2007;13(1):82-90.

52. Tong MJ, Hsu L, Hsien C, et al. A comparison of hepatitis B viral markers of patients in different clinical stages of chronic infection. *Hepatology international.* 2010;4(2):516-522.

53. Mohamadkhani A, Jazii FR, Poustchi H, et al. The role of mutations in core protein of hepatitis B virus in liver fibrosis. *Virology journal.* 2009;6(Journal Article):209.

54. Asim M, Malik A, Sarma MP, et al. Hepatitis B virus BCP, Precore/core, X gene mutations/genotypes and the risk of hepatocellular carcinoma in India. *Journal of medical virology.* 2010;82(7):1115-1125.

55. Zhu Y, Jin Y, Guo X, et al. Comparison study on the complete sequence of hepatitis B virus identifies new mutations in core gene associated with hepatocellular carcinoma. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology.* 2010;19(10):2623-2630.

56. Wu C, Zhang X, Tian Y, et al. Biological significance of amino acid substitutions in hepatitis B surface antigen (HBsAg) for glycosylation, secretion, antigenicity and immunogenicity of HBsAg and hepatitis B virus replication. *The Journal of general virology.* 2010;91(Pt 2):483-492.

57. Cho EY, Kim HC, Choi CS, et al. Nucleotide changes related to hepatocellular carcinoma in the enhancer 1/x-promoter of hepatitis B virus subgenotype C2 in cirrhotic patients. *Cancer science.* 2010;101(8):1905-1912.

58. Durantel D. Fitness and infectivity of drug-resistant and cross-resistant hepatitis B virus mutants: why and how is it studied? *Antivir Ther.* 2010;15(3 Pt B):521-527.

59. Kim do Y, Ahn SH, Lee HW, et al. Clinical course of virologic breakthrough after emergence of YMDD mutations in HBeAg-positive chronic hepatitis B. *Intervirology.* 2008;51(4):293-298.

60. Baldick CJ, Eggers BJ, Fang J, et al. Hepatitis B virus quasispecies susceptibility to entecavir confirms the relationship between genotypic resistance and patient virologic response. *J Hepatol.* Jun 2008;48(6):895-902.

61. Liu F, Chen L, Yu DM, et al. Evolutionary patterns of hepatitis B virus quasispecies under different selective pressures: correlation with antiviral efficacy. *Gut.* Feb 2 2011.

62. Lauring AS, Andino R. Quasispecies theory and the behavior of RNA viruses. *PLoS Pathog.* 2010;6(7):e1001005.

63. Osiowy C, Kaita K, Solar K, Mendoza K. Molecular characterization of hepatitis B virus and a 9-year clinical profile in a patient infected with genotype I. *J Med Virol.* May 2010;82(6):942-948.

64. Torresi J. The virological and clinical significance of mutations in the overlapping envelope and polymerase genes of hepatitis B virus. *Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology.* 2002;25(2):97-106.

65. Solmone M, Vincenti D, Prosperi MC, Bruselles A, Ippolito G, Capobianchi MR. Use of massively parallel ultradeep pyrosequencing to characterize the genetic diversity of hepatitis B virus in drug-resistant and drug-naive patients and to detect minor variants in reverse transcriptase and hepatitis B S antigen. *J Virol.* Feb 2009;83(4):1718-1726.

66. Wang C, Mitsuya Y, Gharizadeh B, Ronaghi M, Shafer RW. Characterization of mutation spectra with ultra-deep pyrosequencing: application to HIV-1 drug resistance. *Genome Res.* Aug 2007;17(8):1195-1201.

67. Cao GW. Clinical relevance and public health significance of hepatitis B virus genomic variations. *World J Gastroenterol.* Dec 14 2009;15(46):5761-5769.

68. Raimondo G, Pollicino T, Romano L, Zanetti AR. A 2010 update on occult hepatitis B infection. *Pathologie-biologie.* 2010;58(4):254-257.

69. Torbenson M, Thomas DL. Occult hepatitis B. *The Lancet infectious diseases.* 2002;2(8):479-486.

70. Raimondo G, Pollicino T, Cacciola I, Squadrito G. Occult hepatitis B virus infection. *Journal of hepatology.* 2007;46(1):160-170.

71. Samuel D, Forns X, Berenguer M, et al. Report of the monothematic EASL conference on liver transplantation for viral hepatitis (Paris, France, January 12-14, 2006). *Journal of hepatology.* 2006;45(1):127-143.

72. Minuk GY, Sun DF, Uhanova J, et al. Occult hepatitis B virus infection in a North American community-based population. *Journal of hepatology.* 2005;42(4):480-485.

73. Schmeltzer P, Sherman KE. Occult hepatitis B: clinical implications and treatment decisions. *Dig Dis Sci.* Dec 2010;55(12):3328-3335.

74. Chemin I, Trepo C. Clinical impact of occult HBV infections. *J Clin Virol.* Dec 2005;34 Suppl 1:S15-21.

75. Liang TJ, Mok KT, Liu SI, et al. Hepatitis B genotype C correlated with poor surgical outcomes for hepatocellular carcinoma. *Journal of the American College of Surgeons.* 2010;211(5):580-586.

76. Kusakabe A, Tanaka Y, Inoue M, et al. A population-based cohort study for the risk factors of HCC among hepatitis B virus mono-infected subjects in Japan. *Journal of gastroenterology.* 2010(Journal Article).

77. Hosaka T, Suzuki F, Kobayashi M, et al. Development of HCC in patients receiving adefovir dipivoxil for lamivudine-resistant hepatitis B virus mutants. *Hepatology research : the official journal of the Japan Society of Hepatology.* 2010;40(2):145-152.

78. Chen CH, Lee CM, Tung WC, et al. Evolution of full-length HBV sequences in chronic hepatitis B patients with sequential lamivudine and adefovir dipivoxil resistance. *Journal of hepatology.* 2010;52(4):478-485.

79. Larke RP, Froese GJ, Devine RD, Petruk MW. Extension of the epidemiology of hepatitis B in circumpolar regions through a comprehensive serologic study in the Northwest Territories of Canada. *Journal of medical virology.* 1987;22(3):269-276.

80. McCance DJ. DNA Viruses: DNA extraction, purification and characterization. In: Mahy BW, Kangro HO, eds. *Virology methods manual.* Vol 1. Toronto, ON: Academic Press Limited; 1996:191-230.

81. Gunther S, Li BC, Miska S, Kruger DH, Meisel H, Will H. A novel method for efficient amplification of whole hepatitis B virus genomes permits rapid functional analysis and reveals deletion mutants in immunosuppressed patients. *Journal of virology.* 1995;69(9):5437-5444.

82. Durantel D, Carrouee-Durantel S, Werle-Lapostolle B, et al. A new strategy for studying in vitro the drug susceptibility of clinical isolates of human hepatitis B virus. *Hepatology (Baltimore, Md.).* 2004;40(4):855-864.

83. Lindh M, Gonzalez JE, Norkrans G, Horal P. Genotyping of hepatitis B virus by restriction pattern analysis of a pre-S amplicon. *Journal of virological methods.* 1998;72(2):163-174.

84. Stuyver L, Van Geyt C, De Gendt S, et al. Line probe assay for monitoring drug resistance in hepatitis B virus-infected patients during antiviral therapy. *Journal of clinical microbiology.* 2000;38(2):702-707.

85. Pollicino T, Squadrito G, Cerenzia G, et al. Hepatitis B virus maintains its pro-oncogenic properties in the case of occult HBV infection. *Gastroenterology.* 2004;126(1):102-110.

86. Lindh M, Furuta Y, Ljunggren KK, Norkrans G, Horal P. Detection of hepatitis B virus precore TAG mutant by an amplification-created restriction site method. *The Journal of infectious diseases.* 1995;171(1):194-197.

87. Takahashi K, Aoyama K, Ohno N, et al. The precore/core promoter mutant (T1762A1764) of hepatitis B virus: clinical significance and an easy method for detection. *The Journal of general virology.* 1995;76 (Pt 12)(Pt 12):3159-3164.

88. Minuk GY, Sun DF, Greenberg R, et al. Occult hepatitis B virus infection in a North American adult hemodialysis patient population. *Hepatology (Baltimore, Md.).* 2004;40(5):1072-1077.

89. Osiowy C, Villeneuve JP, Heathcote EJ, Giles E, Borlang J. Detection of rtN236T and rtA181V/T mutations associated with resistance to adefovir dipivoxil in samples from patients with chronic hepatitis B virus infection by the INNO-LiPA HBV DR line probe assay (version 2). *Journal of clinical microbiology.* 2006;44(6):1994-1997.

90. Zoulim FD, FR), Barraud, Luc (Belleville Sur Saone, FR), Durantel, Sandra (Genas, FR), Lebel-binay, Sophie (Villejuif, FR), Durantel, David (Genas, FR), Trepo, Christian (Bron, FR), Inventor. Method of genotyping and phenotyping hepatitis B viruses resistant to antiviral molecules2006.

91. Datta S, Banerjee A, Chandra PK, et al. Analysis of hepatitis B virus X gene phylogeny, genetic variability and its impact on pathogenesis: implications in Eastern Indian HBV carriers. *Virology.* 2008;382(2):190-198.

92. Chaudhuri V, Tayal R, Nayak B, Acharya SK, Panda SK. Occult hepatitis B virus infection in chronic liver disease: full-length genome and analysis of mutant surface promoter. *Gastroenterology.* 2004;127(5):1356-1371.

93. Rozanov M, Plikat U, Chappey C, Kochergin A, Tatusova T. A web-based genotyping resource for viral sequences. *Nucleic acids research.* 2004;32(Web Server issue):W654-659.

94. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic acids research.* 1997;25(24):4876-4882.

95. Wang HY, Chien MH, Huang HP, et al. Distinct hepatitis B virus dynamics in the immunotolerant and early immunoclearance phases. *Journal of virology.* 2010;84(7):3454-3463.

96. Helgason A, Palsson G, Pedersen HS, et al. mtDNA variation in Inuit populations of Greenland and Canada: migration history and population structure. *Am J Phys Anthropol.* May 2006;130(1):123-134.

97. Wang S, Lewis CM, Jakobsson M, et al. Genetic variation and population structure in native Americans. *PLoS Genet.* Nov 2007;3(11):e185.

98. Minuk G, Uhanova J. Chronic hepatitis B infection in Canada. *The Canadian journal of infectious diseases = Journal canadien des maladies infectieuses.* 2001;12(6):351-356.

99. Minuk GY, Nicolle LE, Postl B, Waggoner JG, Hoofnagle JH. Hepatitis virus infection in an isolated Canadian Inuit (Eskimo) population. *Journal of medical virology.* 1982;10(4):255-264.

100. Baikie M, Ratnam S, Bryant DG, Jong M, Bokhout M. Epidemiologic features of hepatitis B virus infection in northern Labrador. *CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne.* 1989;141(8):791-795.

101. Allain JP. Occult hepatitis B virus infection. *Transfusion clinique et biologique : journal de la Societe francaise de transfusion sanguine.* 2004;11(1):18-25.

102. Brechot C, Thiers V, Kremsdorf D, Nalpas B, Pol S, Paterlini-Brechot P. Persistent hepatitis B virus infection in subjects without hepatitis B surface antigen: clinically significant or purely "occult"? *Hepatology (Baltimore, Md.).* 2001;34(1):194-203.

103. Yoshikawa A, Gotanda Y, Minegishi K, et al. Lengths of hepatitis B viremia and antigenemia in blood donors: preliminary evidence of occult (hepatitis B surface antigen-negative) infection in the acute stage. *Transfusion.* 2007;47(7):1162-1171.

104. Guirgis BS, Abbas RO, Azzazy HM. Hepatitis B virus genotyping: current methods and clinical implications. *Int J Infect Dis.* Nov 2010;14(11):e941-953.

105. Bartholomeusz A, Schaefer S. Hepatitis B virus genotypes: comparison of genotyping methods. *Reviews in medical virology.* 2004;14(1):3-16.

106. Hassan ZK, Hafez MM, Mansor TM, Zekri AR. Occult HBV infection among Egyptian hepatocellular carcinoma patients. *Virol J.* 2011;8:90.

107. Xu L, Wei Y, Chen T, et al. Occult HBV infection in anti-HBs-positive young adults after neonatal HB vaccination. *Vaccine.* Aug 23 2010;28(37):5986-5992.

108. Kumar GT, Kazim SN, Kumar M, et al. Hepatitis B virus genotypes and hepatitis B surface antigen mutations in family contacts of hepatitis B virus infected patients with occult hepatitis B virus infection. *Journal of gastroenterology and hepatology.* 2009;24(4):588-598.

109. Statistics C. *Aboriginal Identity, sex and age groups for the population of Canada*2006.

110. Raimondi S, Maisonneuve P, Bruno S, Mondelli MU. Is response to antiviral treatment influenced by hepatitis B virus genotype? *Journal of hepatology.* 2010;52(3):441-449.

111. Pollicino T, Isgro G, Di Stefano R, et al. Variability of reverse transcriptase and overlapping S gene in hepatitis B virus isolates from untreated and lamivudine-resistant chronic hepatitis B patients. *Antiviral Therapy.* 2009;14(5):649-654.

112. Cacciola I, Pollicino T, Squadrito G, Cerenzia G, Orlando ME, Raimondo G. Occult hepatitis B virus infection in patients with chronic hepatitis C liver disease. *The New England journal of medicine.* 1999;341(1):22-26.

113. Fabrizi F, Messa PG, Lunghi G, et al. Occult hepatitis B virus infection in dialysis patients: a multicentre survey. *Alimentary Pharmacology & Therapeutics.* 2005;21(11):1341-1347.

114. Roche B, Samuel D, Gigou M, et al. De novo and apparent de novo hepatitis B virus infection after liver transplantation. *Journal of hepatology.* 1997;26(3):517-526.

115. Park R. The Dorset-Thule Succession in Arctic North America: Assessing Claims for Culture Contact. *American Antiquity.* 1993;58(2):203-234.

116. Wu CF, Yu MW, Lin CL, et al. Long-term tracking of hepatitis B viral load and the relationship with risk for hepatocellular carcinoma in men. *Carcinogenesis.* 2008;29(1):106-112.

117. Yuen MF, Tanaka Y, Fong DY, et al. Independent risk factors and predictive score for the development of hepatocellular carcinoma in chronic hepatitis B. *Journal of hepatology.* 2009;50(1):80-88.

118. Yeung P, Wong DK, Lai CL, Fung J, Seto WK, Yuen MF. Profile of pre-S deletions in the natural history of chronic hepatitis B infection. *Journal of medical virology.* 2010;82(11):1843-1849.

119. Brind A, Jiang J, Samuel D, et al. Evidence for selection of hepatitis B mutants after liver transplantation through peripheral blood mononuclear cell infection. *Journal of hepatology.* 1997;26(2):228-235.

120. Lizzano RA, Yang B, Clippinger AJ, Bouchard MJ. The C-terminal region of the hepatitis B virus X protein is essential for its stability and function. *Virus research.* 2011;155(1):231-239.

121. Chu CJ, Keeffe EB, Han SH, et al. Prevalence of HBV precore/core promoter variants in the United States. *Hepatology (Baltimore, Md.).* 2003;38(3):619-628.

122. Yang HI, Yeh SH, Chen PJ, et al. Associations between hepatitis B virus genotype and mutants and the risk of hepatocellular carcinoma. *Journal of the National Cancer Institute.* 2008;100(16):1134-1143.

123. Liu S, Zhang H, Gu C, et al. Associations between hepatitis B virus mutations and the risk of hepatocellular carcinoma: a meta-analysis. *J Natl Cancer Inst.* Aug 5 2009;101(15):1066-1082.

124. Guo X, Jin Y, Qian G, Tu H. Sequential accumulation of the mutations in core promoter of hepatitis B virus is associated with the development of hepatocellular carcinoma in Qidong, China. *J Hepatol.* Nov 2008;49(5):718-725.

125. Kajiya Y, Hamasaki K, Nakata K, et al. Full-length sequence and functional analysis of hepatitis B virus genome in a virus carrier: a case report suggesting the impact of pre-S and core promoter mutations on the progression of the disease. *Journal of viral hepatitis.* 2002;9(2):149-156.

126. Milich D, Liang TJ. Exploring the biological basis of hepatitis B e antigen in hepatitis B virus infection. *Hepatology (Baltimore, Md.).* 2003;38(5):1075-1086.

127. Sakamoto T, Tanaka Y, Simonetti J, et al. Classification of hepatitis B virus genotype B into 2 major types based on characterization of a novel subgenotype in Arctic indigenous populations. *The Journal of infectious diseases.* 2007;196(10):1487-1492.

128. Minuk GY, Orr PS, Brown R, Macdonald S, Chaudhary RK, Temple P. Pre-core mutant infections in the Canadian Inuit. *Journal of hepatology.* 2000;33(5):781-784.

129. Wu S, Imazeki F, Kurbanov F, et al. Evolution of hepatitis B genotype C viral quasi-species during hepatitis B e antigen seroconversion. *J Hepatol.* Jan 2011;54(1):19-25.

130. Hannoun C, Norder H, Lindh M. An aberrant genotype revealed in recombinant hepatitis B virus strains from Vietnam. *The Journal of general virology.* 2000;81(Pt 9):2267-2272.

131. Karasawa T, Shirasawa T, Okawa Y, et al. Association between frequency of amino acid changes in core region of hepatitis B virus (HBV) and the presence of precore mutation in Japanese HBV carriers. *Journal of gastroenterology.* 1997;32(5):611-622.

132. Desmond CP, Bartholomeusz A, Gaudieri S, Revill PA, Lewin SR. A systematic review of T-cell epitopes in hepatitis B virus: identification, genotypic variation and relevance to antiviral therapeutics. *Antivir Ther.* 2008;13(2):161-175.

133. Jiang Y, Li W, Yu L, et al. Enhancing the antihepatitis B virus immune response by adefovir dipivoxil and entecavir therapies. *Cell Mol Immunol.* Jan 2010;8(1):75-82.

134. Audsley J, Littlejohn M, Yuen L, et al. HBV mutations in untreated HIV-HBV co-infection using genomic length sequencing. *Virology.* Sep 30 2010;405(2):539-547.

135. Mizokami M, Orito E, Ohba K, Ikeo K, Lau JY, Gojobori T. Constrained evolution with respect to gene overlap of hepatitis B virus. *Journal of Molecular Evolution.* 1997;44 Suppl 1(Journal Article):S83-90.

136. Nakabayashi J, Sasaki A. A mathematical model of the intracellular replication and within host evolution of hepatitis type B virus: Understanding the long time course of chronic hepatitis. *Journal of theoretical biology.* 2011;269(1):318-329.

137. Forbi JC, Vaughan G, Purdy MA, et al. Epidemic history and evolutionary dynamics of hepatitis B virus infection in two remote communities in rural Nigeria. *PLoS One.* 2010;5(7):e11615.

138. Parkinson AJ. The International Polar Year, 2007-2008, an opportunity to focus on infectious diseases in Arctic regions. *Emerg Infect Dis.* Jan 2008;14(1):1-3.

139. Arien KK, Vanham G, Arts EJ. Is HIV-1 evolving to a less virulent form in humans? *Nature reviews.Microbiology.* 2007;5(2):141-151.

140. Zuniga MC. A pox on thee! Manipulation of the host immune system by myxoma virus and implications for viral-host co-adaptation. *Virus research.* 2002;88(1-2):17-33.

141. Knell RJ. Syphilis in renaissance Europe: rapid evolution of an introduced sexually transmitted disease? *Proceedings.Biological sciences / The Royal Society.* 2004;271 Suppl 4(Journal Article):S174-176.

142. Kurbanov F. HBV/D Sequence Comparisons Among Regions; Alaskan HBV/F Paired Sera Analysis. Paper presented at: The International Circumpolar Arctic Viral Hepatitis Working Group; July 11, 2009, 2009; Yellowknife, NT, Canada.

143. Arauz-Ruiz P, Norder H, Visona KA, Magnius LO. Molecular epidemiology of hepatitis B virus in Central America reflected in the genetic variability of the small S gene. *J Infect Dis.* Oct 1997;176(4):851-858.

144. Carman WF, Thomas HC. Implications of genetic variation on the pathogenesis of hepatitis B virus infection. *Archives of virology.Supplementum.* 1993;8(Journal Article):143-154.

145. Rodriguez-Frias F, Buti M, Jardi R, et al. Hepatitis B virus infection: precore mutants and its relation to viral genotypes and core mutations. *Hepatology (Baltimore, Md.).* 1995;22(6):1641-1647.

146. Archary D, Gordon ML, Green TN, Coovadia HM, Goulder PJ, Ndung'u T. HIV-1 subtype C envelope characteristics associated with divergent rates of chronic disease progression. *Retrovirology.* 2010;7:92.

147. Sklan EH, Charuworn P, Pang PS, Glenn JS. Mechanisms of HCV survival in the host. *Nature reviews.Gastroenterology & hepatology.* 2009;6(4):217-227.