Genetic Diversity of Hepatitis B Virus in Indonesia:

Epidemiological and Clinical Significance

Meta Dewi Thedja

ISBN: 9789039358542
Designed by Danang Ispra
Printed by DIC Creative

Genetic Diversity of Hepatitis B Virus in Indonesia:
Epidemiological and Clinical Significance

Genetische Variaties van het Hepatitis B Virus in Indonesië:
Epidemiologische en Klinische Betekenis

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht
op gezag van de rector magnificus, prof. dr. G.J. van der Zwaan,
ingevolge het besluit van het college voor promoties in het openbaar
te verdedigen op vrijdag 19 oktober 2012 des middags te 2.30 uur

door

Meta Dewi Thedja

geboren op 24 april 1962 te Jakarta

Promotoren :

Prof. dr. I.M. Hoepelman
Prof. dr. J. Verhoef
Prof. dr. D.H. Muljono

Leden Beoordelingscommissie :

Prof. dr. P.D. Siersema
Prof. dr. R.A. Coutinho
Prof. dr. P.L. Jansen
Prof. dr. M.D. de Jong
Prof. dr. S. Marzuki
Prof. dr. E.J.H.J. Wierz

Paranimfen:

Korri Elvanita El Khobar
Andisa Dewi

Contents

Chapter 1	General Introduction	1
Chapter 2	Ethnographical structure of hepatitis B virus genotype distribution in Indonesia and discovery of a new subgenotype, B9	25
Chapter 3	Genogeography and immune epitope characteristics of Hepatitis B virus genotype C reveals two distinct types: Asian and Papua Pacific	61
Chapter 4	Occult hepatitis B in blood donors in Indonesia: altered antigenicity of the hepatitis B virus surface protein	93
Chapter 5	Prediction of conformational changes by single mutation in the hepatitis B virus surface antigen (HBsAg) identified in HBsAg-negative blood donors	111
Chapter 6	Viral kinetics in the natural history of chronic hepatitis B in Indonesia	127
Chapter 7	Hepatitis B virus genotype determination using simple restriction fragment length polymorphism (RFLP) for the detection of diagnostic single nucleotide polymorphism (SNPs) in the preS/S region	151
Chapter 8	General discussion, future perspectives, and conclusion	171
	Samenvatting	183
	Ringkasan	189
	Acknowledgement	195
	Curriculum vitae	199
	List of publication	200

CHAPTER 1

General Introduction

General Introduction

Hepatitis B is a necroinflammatory liver disease caused by hepatitis B virus (HBV) which was discovered in 1965 by Baruch Blumberg. HBV infection affects a significant proportion of world population with estimated 240 million people have been diagnosed as chronic hepatitis B (CHB), and are at increasing risk of developing cirrhosis, liver failure and hepatocellular carcinoma [1]. Approximately 75% of those HBV infected-people reside in Asia and Pacific [2] that belong to moderate-to-high hepatitis B endemic region. According to the World Health Organization (WHO), there are 1-2 million deaths/year caused by complications of hepatitis B [3,4] with 50 new cases are diagnosed annually. All these data have increased our alertness and concern of the epidemic hepatitis B in developing countries as well as developed countries.

The disease spectrum of hepatitis B and its complications show high variation, from asymptomatic carrier to chronic hepatitis, hepatic cirrhosis, fulminant hepatitis and hepatocellular carcinoma. However, the precise pathogenetic mechanism responsible for these variable outcomes is poorly defined. Most studies indicate that viral genome characteristics may contribute to HBV persistence [4,5,6]. In addition, host genetic and environmental factors are clearly elicited in HBV pathogenesis [7,8,9,10].

HBV is a partially double-stranded DNA virus with four largely overlapping Open Reading Frames (ORFs) that encode surface, core, x, and polymerase proteins. In evolutionary terms, HBV shows two opposing tendencies: its replication by error-prone reverse transcriptase that leads to elevated mutation rates, and its compact genome with overlap reading frames that relatively limits to genetic variation[11]. These opposite aspects render the substitution rate of HBV to an intermediate level between RNA and DNA viruses [12]. Another implication of this enhanced potential variability is the generation of a quasispecies-like viral population [13], harboring viral mutations that eventually can be selected under particular selection pressure. Therefore, the emergence of HBV with high genome diversity is likely a frequent event that may have public health and clinical implications.

Mutations that occur in surface protein, hepatitis B surface antigen (HBsAg), might be undetectable by diagnostic tools or escape the immune response of host and cause vaccine failure [14,15]. Mutations in precore/core protein have been shown to be associated with more severe complications [16,17]; while mutations in polymerase protein are associated with nucleos(t)ide treatment failure [18,19]. Further, genome characteristics allow classification of HBV into eight genotypes, each with distinct geographical distribution worldwide, and interestingly, following the ethnic background of host [20,21].

Furthermore, different genotypes have been shown to result in different clinical implications and treatment responses [22,23,24,25]. All those data strongly support the importance of HBV genotype information for better clinical management and epidemiological strategies in combating HBV infection.

Indonesia is an archipelago country located in the strategic position of Southeast Asia that bridges the mainland Asia and the Oceania, consisting of around 17,000 islands, which are inhabited by more than 350 ethnic population groups with some in very isolated islands [26]. These Indonesian populations harbor a substantial linguistic diversity and cultural variety in addition to the contrasting phenotype of each ethnic group. This also reflects on the viral genome of HBV with which these populations are infected. A dynamic interaction between viral replication and immune response of host is critical in the pathogenesis and natural course of liver disease. Therefore, the understanding of HBV genetic diversity in various ethnic populations and their inhabitance in the distinct geography is important as a basic knowledge in exploring HBV pathogenesis and epidemiology in a country with high prevalence of CHB such as Indonesia.

I. Hepatitis B virus infection and the liver disease: an overview

Distribution of hepatitis B

The geographical prevalence of HBV infection, defined as the presence of HBsAg, varies widely with three categories of endemicity, low (<2%), intermediate (2-8%) and high (>8%) prevalence. The rate of HBsAg carriage in the general population shows a wide range, from 2% to 20%, with special notification that it is prevalent in Asia including Southeast Asia and Pacific [27]. In these areas, about 70-90% of the populations become HBV-infected before the age of 40, and 8-20% of the people are carriers [28]. The prevalence of chronic HBV infection in the Asia and Pacific is among the highest in the world with different rates in some countries [29,30]. The HBsAg prevalence is around 6-9% in Indonesia and northern China, and higher (>10%) in Taiwan, southern China, Korea, Philippines, Melanesia, Micronesia, and Polynesia. In these highly endemic countries, around 70-95% of the populations have shown both the past and present serological evidences of HBV infection [30].

Most developed countries are classified into low endemic HBV infection areas, including North America, Western and Northern Europe, Australia, and parts of South America. However, immigrants from different ethnicities especially from highly endemic countries might contribute to the local prevalence that could change the endemicity status of the country [31]. The rest of the world falls into the intermediate range of HBV prevalence with

2-8% of populations being chronic HBV carriers [28]. In addition, globally, approximately 45% of the global populations live in areas with high HBV prevalence [32].

Transmission and vaccination of hepatitis B

Chronic carriage of HBV is defined as the presence of HBsAg, persistent > 6 months after infection with or without elevation of serum alanine aminotransferase [33]. Chronic hepatitis B constitutes a health problem since the carriers are reservoirs of infectious agents that can spread horizontally or vertically to other individuals by percutaneous or mucosal exposure to infected blood or other body fluids. Perinatal transmission, household contact, sexual contact, blood transfusion, semen and unsterilized injection or needle-sharing practices are known as common routes of HBV transmission. The route of HBV transmission and the age of infection have important clinical implications. In high-prevalence areas, perinatal/vertical transmission or horizontal infections in early childhood are the most common routes of HBV transmission, such as in Asia, Africa, Pacific Islands, or in Arctic. In contrast, in low endemic countries, hepatitis B is commonly acquired horizontally by exposure to HBV contaminated tools or through risky behavior, such as unprotected sexual contact or sharing syringes of personal equipments with HBsAg positive individuals [31,34,35].

Up to 95% infants from mothers HBsAg- and Hepatitis B *e* antigen (HBeAg)-positive are at risk of HBV acquisition; about 70-90% of these perinatally infected-newborns will develop chronic, usually lifelong hepatitis B infection, if no immunoprophylaxis is given [27,35]. The risk of perinatal infection among infants born to HBsAg-positive and HBeAg-negative mother's ranges from 10-40% with 40-70% of these infants remaining chronically infected [36]. On the other hand, the risk of developing chronic infection from horizontal transmission during adulthood is less than 5% [35]. In areas of low endemicity, perinatal and early childhood transmission may also account for more than one third of chronic infections [35]. Thus, it is crucial to prevent perinatal transmission of HBV via the identification and treatment of HBsAg in pregnant women, and the administration of immunoprophylaxis to their newborns. In an attempt to reduce HBV vertical transmission, CDC has recommended prenatal HBsAg-screening for all pregnant women since 1988.

In addition to prenatal screening of all pregnant women for HBsAg, vaccination of infants is a key strategy for the prevention of HBV infection spreading. In 1992, the World Health Assembly passed a resolution calling all WHO Member States to, where feasible, integrate the hepatitis B vaccination into national immunization programs. In the same year, WHO set a goal for all countries to introduce the hepatitis B vaccine into national routine infant immunization programs by 1997 [37]. The implementation of universal infant

immunization since three decades ago has reduced HBV infection rate in highly endemic-countries. The inclusion of hepatitis B vaccine into global infant immunization program have contributed to prevent >80% of HBV-related deaths [38]. However, despite the presence of hepatitis B vaccine, new HBV infection continues to occur via vertical and horizontal transmission routes [35,39]. Recently, antiviral treatment during the last trimester of pregnancy in HBsAg-positive women has been recommended by three international bodies (EASL, APASL and AASLD) [40,41,42]. It is expected that this strategy is followed by public health authorities to reduce the acquisition of HBV infection through mother-to-child transmission in wider communities.

Natural history of HBV infection

The outcome of hepatitis B greatly varies among individuals and is associated with a diverse clinical spectrum of liver damage, ranging from acute or fulminant hepatitis to various forms of chronic liver diseases, including inactive carrier state, chronic hepatitis, cirrhosis, and hepatocellular carcinoma (HCC). Two important points that need to be underlined in the natural outcome of acute HBV infection are age and immune competence at the time of infection [27]. In children and infants, initial HBV infection is typically subclinical or asymptomatic, and most of acute cases proceed to chronic infection. Generally, acquisition of initial HBV infection during the adulthood period is symptomatic with icteric clinical appearance (Figure 1A), in which and fulminant hepatitis occurs in 0.1-0.5% of acute HBV infection cases.

In chronic HBV infection, there are four phases of infection: immune tolerance (IT), immune clearance (IC), low or non-replicative (LR), and reactivation phases. In the IT phase, immune activity against the virus is low, and viral replication is high with detectable HBV DNA in serum (Figure 1B). In this phase, serum ALT level is within normal limits and generally, patients are without any symptoms of disease, and on biopsy the liver shows minimal inflammatory activity [27]. In perinatal infection, this phase can last for the first two decades of life with low rate of spontaneous HBsAg clearance [43].

In the second (IC) phase of infection, clinical manifestation occurs as a result of immune mediated destruction of HBV infected-hepatocytes, leading to elevated serum ALT or liver enzymes, decreased HBV DNA level, and the presence of hepatitis B e antigen (HBeAg). The duration of this second phase varies from months to years [27]. The third phase of infection, or more commonly referred as 'inactive carrier' state, is characterized by normalization of serum ALT, low or undetectable HBV DNA, HBeAg seroconversion and histopathologal improvement, followed by sustained clinical remission. In this phase, more rapid progression is seen due to cessation of viral replication and regression of

fibrosis may occur months to years after seroconversion of HBeAg. However, this inactive state can revert back to IC phase with HBeAg seropositivity, or maintain HBeAg seronegativity and develop into HBeAg-negative hepatitis (ENH) [42,45], which is also categorized as a separate phase of the natural history of CHB [41]. In this ENH phase, serum HBV DNA and ALT levels may increase due to activation of viral replication and inflammatory reaction, followed by progression to fibrosis and liver decompensation.

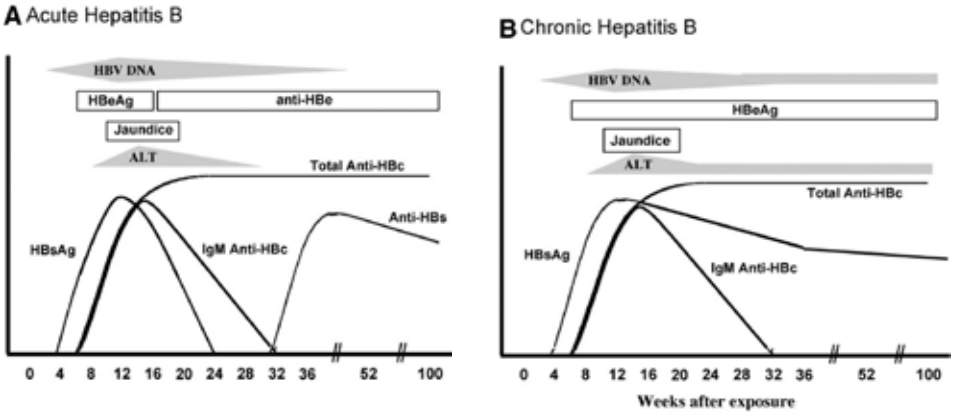


Figure 1. Serological and molecular profiles in acute **(A)** and chronic hepatitis B **(B)** [44].

II. Molecular virology of hepatitis B virus

Family and morphology of hepatitis B virus

HBV belongs to hepadnaviridae family, a family of enveloped viruses with partially double-stranded relaxed circular (rc) DNA genome of 3.2 kb in length, and enveloped by lipid bilayers bearing three different surface proteins. Electron microscopic examination reveals three types of virus-associated particles: (1) HBV virion with 42 nm in diameter that entitles a complete virus including the genome and all proteins inside, known as Dane particle. The outer envelope formed by the HBsAg that surrounds the inner nucleocapsid is made up by the hepatitis B core antigen (HBcAg); (2) spherical particles of around 22 nm in diameter that are in 104 to 106 fold excess of HBV virions; (3) filamentous particles of approximately 20-22 nm in diameter with various lengths. These two latter particles consist of surface protein (HBsAg) without genome and other proteins, and thus, are not as infectious as the virion (Figure 2A). Usually, the non-infectious particles are produced in a 1,000 to 1,000,000-fold excess over virions [46].

Genome of hepatitis B virus

HBV genome is a circular DNA molecule that consists of four Open Reading Frames (ORFs) overlapping genes encoding the viral envelope, nucleocapsid, polymerase and X protein (Figure 2B). The S (surface) gene with around 1.200 bp length has three start codons which are divided into three regions, preS1, preS2, and S regions, that encode the hepatitis B surface protein known as HBsAg. The HBV envelope protein is composed of three HBsAg forms; the so-called large HBsAg (encoded by preS1/preS2/S), middle HBsAg (encoded by preS2/S), and small HBsAg (encoded by S). The small HBsAg is 226 amino acids length, while the middle HBsAg contains 55 amino acids of the preS2 region, and the large by 108-119 amino acids of the preS1 region. In addition, the preS1 region consists of hepatocyte receptor-binding site within residue 21-47, while preS/S contains B-cell and T-cell immune epitopes [47]. The preS2 region, contains only 165 bp and shows the highest nucleotide polymorphism among the S regions. Furthermore, HBsAg is important as the main target for immune response of the host and thus, as the target for diagnostic tool.

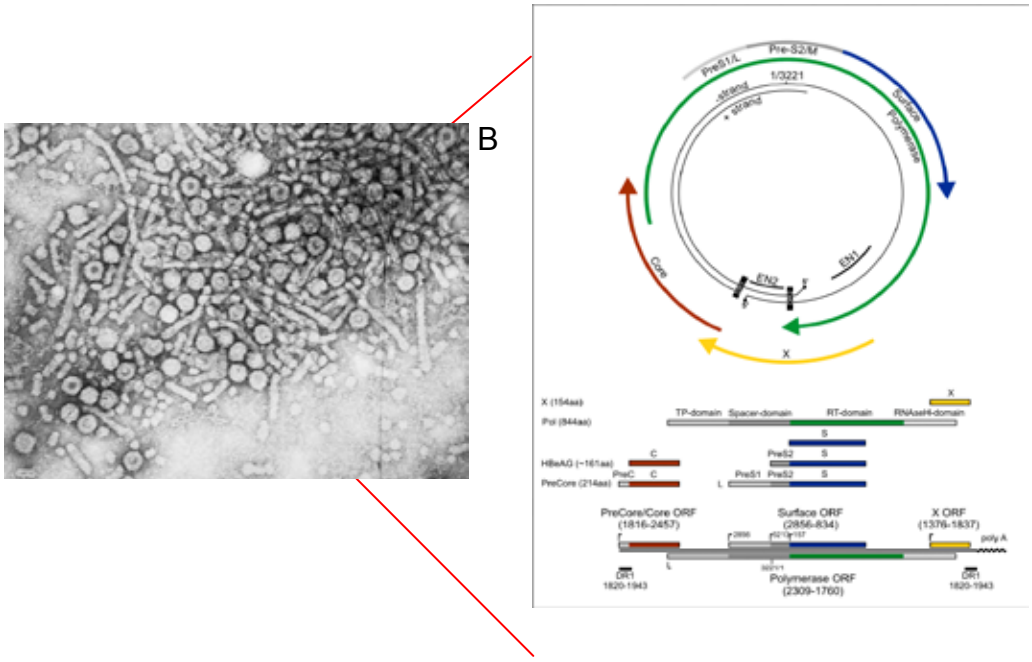


Figure 2. The electron micrograph shows the three forms of HBsAg, the infectious viral (Dane particle) and non-infectious particle **(A)**. HBV genome organization and the translated proteins **(B)** [33, 48].

The precore/core gene is translated into a precore polypeptide of 183 amino acids, which is then modified into a soluble protein HBeAg that serves as a marker of active viral replication, and the nucleocapsid protein, HBcAg. The core protein has a critical role in the interactions between HBV and the immune response of the host and this shapes the course of the disease. HBcAg contains numerous of T and B-cells immune recognition epitopes, with residues 18-27 as the dominant Human Leukocyte Antigen (HLA)-A2-restricted CTL epitope that can elicit a vigorous CTL response in up to 90% of HLA-A2-positif patients with acute self limited HBV infection [7,50,51]. Sequence variation within this immune recognition epitopes has been reported and evaluated for therapeutic studies [52,53,54].

The long polymerase gene encodes DNA polymerase which also serves as reverse transcriptase function, since replication requires RNA intermediates. The X gene encodes HBx protein that transactivates transcriptional promoters, and this protein may play partly in development of hepatocellular carcinoma. These four genes overlap each other and some overlap almost completely the entire genes.

Life cycle and replication of hepatitis B virus

Since HBV is a DNA virus, it replicates through an RNA intermediate in host hepatocyte using its own reverse transcriptase [55]. The rate of virion production is as high as 10^{10}-10^{11} virion/day and an estimated mean of half-life of serum HBV DNA is around 1-2 days. The inefficiency of its genome is the lack of proofreading mechanism of HBV polymerase, which accounts the mutation rate to 1.4-3.2 X 10^{-5} nucleotide substitutions per site per year [56]. The life cycle of HBV is characterized by the synthesis of the 3.2-kb partially double-stranded relaxed circular DNA (rcDNA) following reverse transcription of the 3.5 kb-pregenome RNA [57,58].

Three early steps of HBV life cycle: attachment, penetration and uncoating

The mechanism of these early steps is not well understood due to the limitation of cell lines that are susceptible to hepadnavirus infection [57]. In contrast, the next steps in HBV replication including RNA-directed DNA synthesis have been well characterized [57,59].

The first step is to repair the relaxed circular DNA (rcDNA) viral to covalently closed circular DNA (cccDNA) in which HBV persists as a pool in host cells and serves as a reservoir for virus replication (Figure 4). The formation of cccDNA, which is responsible for the template of viral replication, indicates a successful initiation of infection and can be detected in the liver within 24 hours following virus inoculation [61]. The viral genome is

organized into 4 transcription units, and controlled by 4 independent promoters and a single common polyadenylation signal, yielding 4 extensively overlapping viral RNAs; pre-S1 mRNA, preS/S mRNA, precore mRNA, and pregenomic RNA (pgRNA) that are exported into the cytoplasm where the viral proteins are translated and viral particle assembly and genome replication occur [62]. The pgRNA is subsequently encapsidated in the cell cytoplasm together with a molecule of HBV DNA polymerase. This DNA polymerase has reverse transcriptase function that catalyzes the synthesis of the negative strand genomic RNA. The pgRNA is then gradually degraded by the RNase H activity of the polymerase in the nucleocapsid [63]. Further, a positive DNA strand is synthesized by the polymerase using negative strand as the template. Some new generated nucleocapsid can be re-exported into nucleus to yield additional cccDNA, and some into endoplasmic reticulum to form the complete virion and then they are released to pericellular space by exocytosis.

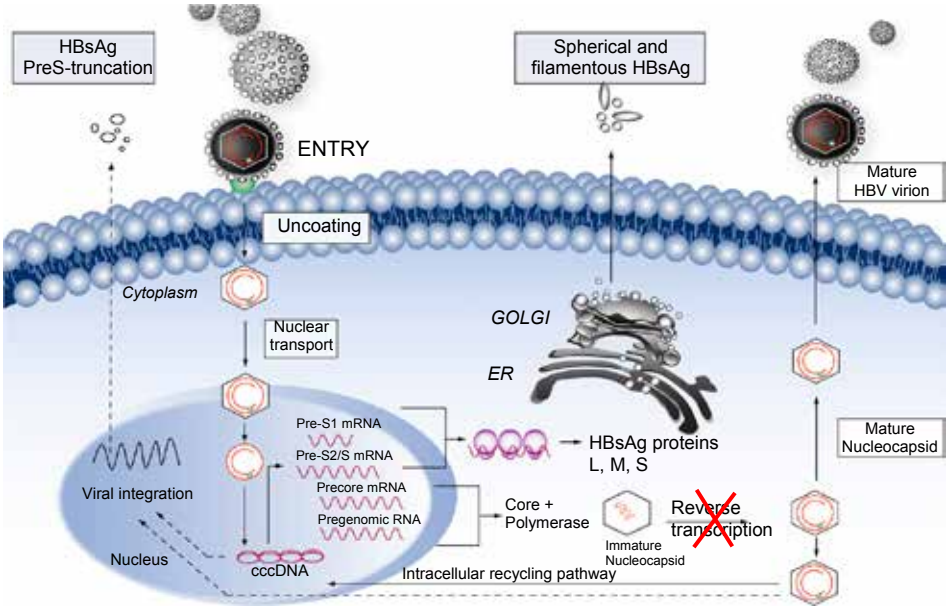


Figure 3. Three highlighted steps of HBV life cycle: (i) the formation of cccDNA as the transcriptional template viral genes; (ii) translation mRNA into proteins; (iii) pgRNA used for progeny genome production [60].

HBV genetic diversity, genotype/subgenotype and HBsAg subtype

HBV genome is highly constrained due to its small size of around 3200 bp lengths and extensive overlapping of the four open reading frames. However, due to the lack of proofreading capacity during its reverse transcription, high nucleotide variability occurs across the genome, including in the *'a'* determinant of HBsAg, which is the most conserved region of the genome. In addition, frequent recombination at both inter- and intra-genotype level can rapidly increase the genome diversity. HBV has been classified into eight genotypes, A to H, differing from each other by sequence divergence exceeding 8% of the complete genome, or more than 4% of the complete surface genome [10,64]. Recently, two genotypes (I and J) were found in Asia and tentatively proposed as new genotypes [65,66]. Further, with the exception of genotype E, G, and H, HBV genotypes have been categorized into several subgenotypes based on sequence divergence of 4–8% of the entire genome [10]. Genotypes B and C, which are predominant in the East and Southeast Asia, can be divided into subgenotypes. Genotype B has been classified into nine subgenotypes B1 – B9, and genotype C into sixteen subgenotypes C1 – C16 [21,67,68,69,70,71,72].

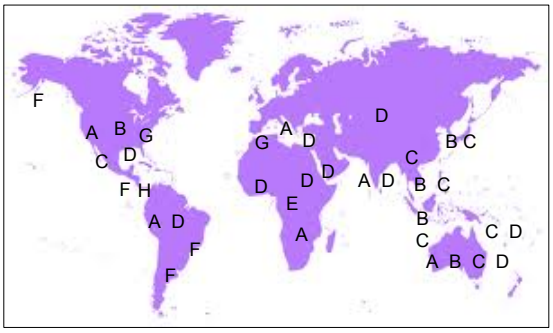


Figure 4. Distribution of HBV genotype worldwide, with genotypes B and C predominant in Asia and Pacific.

HBV genotypes and subgenotypes have distinct geographical distribution and present demographic characteristics (Figure 5). Together with genotype D, genotype A with its five subgenotypes have global distribution found in Europe, Africa, Mediterranean, Middle East region, India, and America. Genotype B and C with their subgenotypes prevail in the East and Southeast Asia. Interestingly, among the nine subgenotypes of genotype B and the sixteen subgenotypes of genotype C, some are found across the Indonesian archipelago: B2, B3, B5, B8, and B9 of genotype B, and C1, C2, C5, C6, C8, C10 to C16 of genotype C. Further, genotype E is mostly found in western Africa; F in

indigenous population in the Central and South America; G in France and United States; and H in Mexico and Central America [10,73,74,75].

There is growing evidence in support of the role of HBV genotype in the activity and disease progression. Some studies performed in Asia have shown a correlation between HBV genotype and severity of disease. In Japan and China, in chronic hepatitis B patients, HBV genotype C is more associated with higher rate of HBeAg carriers, lower rate of spontaneous HBeAg seroconversion, higher levels of HBV DNA, higher histological activities, and higher proportion of patients developing cirrhosis and HCC [22,23,24]. Therefore, the concept that HBV genotype may influence the course of the disease and clinical management would become a major concern for scientists and clinicians.

In addition to genotypes, based on some antigenic determinants of HBsAg, HBV has been characterized serologically into nine subtypes designated as *adw2*, *adw4, adrq+, adrq-, ayw1, ayw2, ayw3, ayw4* and *ayr* [76,77,78]. Together with genotypes, HBV subtypes are distributed geographically specific and follows the ethnic background of the host. In Indonesia, subtype *adw* and genotype B are prominent in the western islands; subtype *adrq+*/*adrq-* and genotype C in eastern islands; and, genotype B and C in between the western and eastern islands of the Indonesian archipelago [21,69,79,80]. Recently, another subdeterminant *q* was found in Papua designated as *adrq-indeterminate* since they showed unusual A159/A177 pair in the positions used for subdeterminant *q* determination [68].

Recombination

Recombination in hepadnavirus was first demonstrated following in vivo DNA transfections using duck HBV [81]. In an endemic area, an infection with more than one genotype often results in genetic recombination among viruses, and this recombination together with the lack of proofreading have increased the diversity of HBV characteristics. In addition, recombination between genotypes occurs in geographical regions where a number of genotypes co-circulate, and provides a mechanism of variation within individuals and populations [10]. Recombination has been reported between genotypes A and D in Africa and India [82], and between genotypes B and C in Asia with genotype B (B1) in Japan showing no recombination within its genome [83]. As other regions with co-existing of two or more genotypes, most of genotype B in Indonesia has been reported to have recombination with genotype C [70]. The presence of HBV strains resulted from recombination between genotypes/subgenotypes could bring

consequences in regard to increasing the diversity of HBV genetic characteristics that may affect the clinical outcome, prognosis, and response to treatment.

III. HBV mutants: clinical and public health significance

The error-prone enzyme reverse transcriptase lacking 3'-5' exonuclease proofreading capacity results in large number of nucleotide substitutions across the genome during replication. The constraints genome and the overlapping of the four open reading frames can easily increase the diversity of the genome. The misincorporation rate has been estimated to be of the order of 10^{10} incorrect nucleotide incorporation per day [84]. As result, HBV circulates as a complex mixture of genetically distinct variants in an infected individual or namely a '*quasispecies*'.

Association of mutation in preS/S region and immune escape, occult hepatitis B, and development of HCC

The Major Hydrophilic Region (MHR) is an important region of the surface protein spanning residue 103 to 173, and exposed to the surface of the particles [85]. This region has 3 major loops held together with disulphide bonds, and the dominant epitope cluster of HBsAg is the *'a'* determinant the region between residue 122 and 149 of the loop 2 and 3 [15,85]. Most anti-HBs in sera from vaccinees binds to amino acid 139 to 147 [15], and variation within these residues can affect the binding sites of antibody and may cause loss of the conformational epitope and altered antigenicity of HBsAg, resulting in HBsAg undetectability, or permitting escape variants to evade virus clearance. Therefore, amino acid variation of HBV genome particularly in preS/S region potentially has clinical and public health implication.

A variety of mutations have been identified in the HBsAg protein resulting in difference antigen recognition and immune response [86,87]. Further, variants within this region have been shown to be associated with the failure of vaccination and the occurrence of escape mutant undetectable HBsAg by current commercial assays [14,88,89,90,91,92,93]. The presence of escape mutant is potential hazardous for the safety of blood donation with regard to introducing occult hepatitis B to the recipient [94,95,96,97,98]. Occult hepatitis B is defined by the presence of HBV DNA in individuals testing HBsAg negative by currently available assays [99]. A G145R mutation is by far the most common immune escape mutant. Other mutations have also been reported and garnered attention; one mutation, T143M, identified in blood donors with occult hepatitis B in Indonesia has garnered attention as it shown to alter HBsAg predicted antigenicity [87,95]. Beside the mutation

affecting HBsAg, variation or deletion in preS2 region has been associated with an increased risk of hepatocellular carcinoma [100,101,102].

Association of mutation in precore/core region and hepatocellular carcinoma

HBV mutations in the precore/core region have been investigated extensively particularly in Asia, the endemic region for HBV infection. In addition to viral load and other risk of HBV chronic infection, two HBV mutants in precore/core region have been suggested to be linked with increasing risk of hepatocellular carcinoma and worsening of chronic liver failure. The predominant mutation in the precore region is G1896A [25,60,16,101] in which creating a stop codon at residue 28 that lead to premature termination of precore/core protein translation and prevents HBeAg production. This mutation is frequently detected in patients with HBeAg-negative chronic hepatitis B and in patients with fulminant hepatic failure [103,104].

The other mutation is the double mutation A1762T/G1764A in the basal core promoter (BCP), which decreases HBeAg production by up to 70% through suppressing the transcription of precore mRNA but enhances viral genome replication in vitro [105]. In the in-vitro transfection studies, these double mutations have been associated with high HBV production with an increased risk of hepatocellular carcinoma [106,107].

Association of mutation in polymerase region and drug resistance

Treatment of chronic hepatitis B remains a clinical and life science challenge due to resistance problems. Some nucleos(t)ide analogues have been approved for the treatment of chronic HBV infection. These nucleos(t)ide analogues suppress HBV replication, induce normalization of serum transaminase level and improve liver histology, thus, preventing the rapid progression of the disease. However, prolonged therapy may result in emergence of drug-resistant HBV [34,108,109].

Most lamivudine resistance HBV shows substitution (rtM204V or rtM204I) in the YMDD (thyrosine-methionine-aspartate-aspartate) motif in the subdomain C of reverse transcriptase domain of polymerase protein [110], whereas rtL180M variant usually appears as compensatory mutation in conjunction with rtM204V/I variant and augment the resistance [108,111]. HBV harboring a N236T and/or A181V variants of the D and B subdomains of reverse transcriptase domain of polymerase, respectively, have been associated with adefovir resistance [112]. Entecavir resistance mutations occurred with combination of variants I169T and M250V, or T184G and S202I [113]. A M204I mutation in the viral polymerase protein has been associated with telbivudine resistance-HBV [109].

IV. Host genetic variability and the outcome of chronic hepatitis B

The outcome of chronic hepatitis B greatly varies, from asymptomatic carrier to cirrhosis, fulminant hepatic failure and hepatocellular carcinoma. Although the pathogenetic mechanism is not well understood, previous studies on epidemiological investigations of human suggests that there is a strong genetic variability to affect the individual susceptibility to infectious pathogen [8,114]. Investigation by Genome-wide Association Study *(*GWAS) approach on some candidate genes of HLA which play critical role in the host immune response to viral infection has been discovered.

Human genetic susceptibility to HBV infection: HLA class I and class II alleles

Most human genetic studies on immune responses to HBV infection focused on HLA analysis. The genes for HLA class I, HLA-A, -B, and –C, and HLA class II, HLA-DRB1, -DQA1, -DQB1, -DPA1, and –DPB1, are located on the short arm of chromosome 6 [115]. HLA class I molecules are responsible for presenting viral peptide-epitope to CD4 T cell, and HLA class II molecules for controlling CD8 T cell function, both of these cells are able to interact directly with infected hepatocytes [8].

HLA class I A*0301 allele and HLA class II DRB1*1301/2 allele are found to be associated with HBV clearance in Gambia and Korea patients [116,117,118]. HLA class I B*08 and B*44, and HLA class II DR7 (DRB1*0701), DR3 (DRB1*0301), DQA1*0301, and DQA1*0501 are found associated with HBV-persistent infection [118,119,120,121]. With regard to the host genetic varieties and their different outcome, variants or polymorphism in HLA class I and class II molecules are of interested to investigate in-depth in particular for Indonesian populations that have very diverse ethnic background.

Aim of the study

To understand the association of genetic diversity of hepatitis B virus in Indonesia and the ethno-geographical background of host in relation to epidemiology medicine and clinical significance.

References

1. World Health Organization (2012). Hepatitis B Fact Sheet No. 204. http://www.who.int/mediacentre/factsheets/fs204/en/.
2. Liaw, Y.F., Leung, N., Kao, J.H., Piratvisuth, T., Gane, E., Han, K.H., *et al*. Asian–Pacific consensus statement on the management of chronic hepatitis B: a 2008 update. Hepatol. Int. 2008, 2: 263–283.
3. Lavanchy, D. Epidemiology (p181-192). In Section IV: Hepatitis B Virus and other Hepadnaviridae. In Viral Hepatitis. (2005). Editors: Thomas, H., Lemon, S., Zuckerman, A. Blackwell Publishing Ltd., Oxford, UK.
4. Huang, Y., Lok, A. Viral factors and outcomes of chronic hepatitis B. Am. J. Gastroenterol. 2011, 106: 93-95.
5. Szmaragd, C., Balloux, F. The population genomics of hepatitis B virus. Molecular. Ecology. 2007, 16: 4747-4758.
6. McMahon, B.J. The influence of hepatitis B virus genotype and subgenotype on the natural history of chronic hepatitis B. Hepatol. Int. 2009, 3: 334-342.
7. Bertoletti, A,. Southwood, S., Chestnut, R., Sette, A., Falco, M., Ferrara, G.B., *et al*. Molecular features of the hepatitis B virus nucleocapsid T-cell epitope 18-27: interaction with HLA and T-cell receptor. Hepatology. 1997, 26: 1027-1034.
8. Thursz, M., Yee, L., Khakoo, S. Understanding the host genetics of chronic hepatitis B and C. Semin. Liver. Dis. 2011, 31: 115-127.
9. Jazayeri, M.S., Basuni, A.A., Cooksley, G., Locarnini, S., Carman, W.F. Hepatitis B virus genotypes, core gene variability and ethnicity in the Pacific region. J. Hepatol. 2004, 41:139–146.
10. Norder, H., Couroucé, A.M., Coursaget, P., Echevarria, J.M., Lee, S.D., Mushahwar, I.K., *et al*. Genetic Diversity of Hepatitis B Virus Strains Derived Worldwide: Genotypes, Subgenotypes, and HBsAg Subtypes. Intervirology. 2004, 47: 289-309.
11. Kann, M., Gerlich, W.H. Structure and molecular virology. In: Thomas HC, Lemon S, Zuckerman AJ, editors. Viral Hepatitis. 3rd ed. Malden: Blackwell; 2005. p. 149–180.
12. Kidd-Ljunggren, K., Miyakawa, Y., Kidd, A.H. Genetic variability in hepatitis B viruses. J. Gen. Virol. 2002, 83: 1267–1280.
13. Gunther, S., Fischer, L., Pult, I., Sterneck, M., Will, H. Naturally occurring variants of hepatitis B virus. Adv. Virus Res. 1999, 52: 25–137.
14. Okamoto, H., Yano, K., Nosaki, Y., Matsui, A., Miyaki, H., Yamamoto, K., *et al*. Mutations within the S gene of hepatitis B virus transmitted from mothers to babies immunized with hepatitis B immune globulin and vaccine. Pediatr. Res. 1992, 32: 264-268.
15. Carman, W.F., van Deursen, F.J., Mimms, L.T., Hardie, D., Coppola, R., Decker, R., *et al*. The prevalence of surface antigen variants of hepatitis B virus in Papua New Guinea, South Africa, and Sardinia. Hepatology. 1997, 26: 1658-1666.

16. Ren, X., Xu, Z., Liu, Y., Li, X., Bai, S., Ding, N., *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. J. Viral. Hepat. 2010, 17: 887-895.

17. Chan, H.L.Y., Thompson, A. , Martinot-Peignoux, M., Piratvisuth, T., Cornberg, M., Brunetto, M.R., *et al*. Hepatitis B surface quantification: why and how to use it in 2011 – a core group report. J. Hepatology. 2011, 55: 1121-1131.

18. Lai, C.L., Ratziu, V., Yuen, M.F., Poynard, T. Viral hepatitis B. Lancet. 2003, 362: 2089-2094.

19. Lee, Y.S., Suh, D.J., Lim, Y.S., Jung, S.W., Kim, K.M., Lee, H.C., *et al*. Increased risk of adefovir resistance in patients with lamivudine-resistant chronic hepatitis B after 48 weeks of adefovir dipifoxil monotherapy. Hepatology. 2006, 43: 1385-1391.

20. Jazayeri, M.S., Carman, W.F. Virus escape CTL or B cell epitopes ? Hepat. Monthly. 2005, 5: 133-136.

21. Thedja, M.D., Muljono, D.H., Nurainy, N., Sukowati, C.H.C., Verhoef, J., Marzuki, S. Ethnogeographical structure of hepatitis B virus genotype distribution in Indonesia and discovery of a new subgenotype, B9. Arch. Virol. 2011; 855-868.

22. Kao, J.H., Chen, P.J., Lai, M.Y., Chen, D.H. Hepatitis B genotypes correlate with clinical outcome in patients with chronic hepatitis B. Gastroenterol. 2000, 118: 554-559.

23. Orito, E., Mizokami, M., Sakugawa, H., Michitaka, K., Ishikawa, K., Ichida, T., *et al*. A Case-control study for clinical and molecular biological differences between Hepatitis B Viruses of genotypes B and C. Hepatology. 2001, 33: 218-223.

24. Zeng, G., Wang, Z., Wen, S., Jiang, J., Wang, L., Cheng, J., *et al*. Geographic distribution, virologic and clinical characteristics of hepatitis B virus genotype in China. J. Viral. Hepat. 2005, 12: 609-617.

25. Yang, H.I., Yeh, S.H., Chen, P.J., Iloeje, U.H., Jen, C.L., Su, J., *et al*. Association between hepatitis B virus genotype and mutants and the risk of hepatocellular carcinoma. J. Natl. Cancer. Inst. 2008, 100: 1134-1143.

26. Sugiono, D. The mapping of languages in Indonesia. Language Center Press 2008, Jakarta, p 7.

27. Tran, T.T., Martin, P. Hepatitis B: epidemiology and natural history. Clin. Liver. Dis. 2004, 8: 255-266.

28. WHO sheet. WHO/CDS/CSR/LYO/2002.2:Hepatitis B.

29. El-Serag, H.B., Rudolph, K.L. Hepatocellular carcinoma: epidemiology and molecular carcinogenesis. Gastroenterology. 2007,132: 2557-2576.

30. Hou, J., Liu, Z., Gu, F. Epidemiology and prevention of Hepatitis B Virus infection. Int. J. Med. Sci. 2005, 2(1): 50-70.

31. McMahon, B.J. Epidemiology and natural history of hepatitis B. Semin. Liver. Dis. 2005, 25(Suppl 1): 3–8.

32. Mahoney, F.J. Update on diagnosis, management, and prevention of hepatitis B virus infection. Clin. Microbiol. Rev. 1999,12:351-366.

33. Lee, W.M. Hepatitis B virus infection. NEJM 2007, 337: 1733-1745.

34. Lai, C.L., Dienstag, J., Schiff, E., Leung, N.W.Y., Atkins, M., Hunt, C., *et al*. Revalence and clinical correlates of YMDD variants during lamivudine therapy for patients with chronic hepatitis B. Clin. Infect. Dis. 2003; 36: 687–696.

35. Zhang, F., Zhang, Y. Perinatal transmission of hepatitis B virus: could hospital doing more? Expert. Rev. Anti. Infect. Ther. 2010, 8(7): 735-738.

36. Hyams, K.C. Risks of Chronicity Following Acute Hepatitis B Virus Infection: A Review. Clin. Infect. Dis. 1995, 20(4): 992-1000.

37. Centers for Disease Control and Prevention (CDC). Global progress toward universal childhood hepatitis B vaccination, 2003. MMWR Morb. Mortal. Weekly. Rep. 2003, 56: 868-870.

38. Goldstein, S.T., Zhou, F., Hadler, S.C., Bell, B.P., Mast, E.M., Margolis, H.S. International. J. Epid. 2005, 34: 1329-1339.

39. Shen, T., Yan, X.M., Zou, Y.L., Gao, J.M., Dong, H. Virologic characteristics of hepatitis B virus in patients infected via maternal-fetal transmission. World. J. Gastroenterol. 2008, 14(37): 5674-5682.

40. Lok, A.S.F., McMahon, B.J. Chronic hepatitis B: update 2009. Hepatology 2009; 50: 1-35.

41. EASL (2012). EASL clinical practice guidelines: management of chronic hepatitis B virus infection. J Hepatology 2012; 57: 167-185.

42. Liaw, Y.F., Kao, J.H., Piratvisuth, T., Chan, H.L.Y., Chien, R.N., Liu, J.C. *et al*. Hepatology Int 2012. DOI 10.1007/s12072-012-9365-4.

43. Chang, M.H. The natural history of hepatitis B virus infection in children. J. Gastroenterol. Hepatol. 2000,15 (Suppl): E16-19.

44. Liang, T.J. Hepatitis B: the virus and disease. Hepatology. 2009; 49: S13-S21.

45. Liaw, Y.F., Chu, C.M. Hepatitis B virus infection. Lancet 2009; 373: 582-592.

46. Grimm, D., Thimme, R., Blum, H.E. HBV life cycle and novel drug targets. Hepatol. Int. 2011, 5: 644-653.

47. Millich, D.R., Thornton, G.B., Neurath, A.R,, Kent, S.B., Michel, M.L., Tiollais, P., *et al*. Enhanced immunogenicity of the PreS region of hepatitis B surface antigen. Science. 1985, 228: 1195-1199.

48. Lüsebrink, J., Schildgen, V., Schildgen, O. HBV – Virology. In: Hepatology. A clinical text book. Mauss, S, Berg, T., Rockstroh, J., Sarrazin, C., Wedemeyer, H. (eds). Hepatology. 2009, p 55-74. Flying Publisher, Dusseldorf, Germany.

49. Liu, H.G., Chen, W.W., Fan, Z.P., Yang, H.Y., Shi, M., Zhang, Z., *et al*. The high prevalence of the 127 mutant HBcAg 18-27 epitope in Chinese HBV-infected patients and its cross-reactivity with the V27 prototype epitope. Clin. Immunol. 2007, 125: 337-345.

50. Bertoletti, A., Chisari, F.V., Penna, A., Guilhot, S., Galati, L., Missale, G., *et al*. Definition of a minimal optimal cytotoxic T-cell epitope within the hepatitis B virus nucleocapsid protein. J. Virol. 1993, 67: 2376-2380.

51. Bertoni, R., Sidney, J., Fowler, P., Chesnut, R.W., Chisari, F.V., Sette, A. Human histocompatibility leukocyte antigen-binding supermotifs predict broadly cross-reactive cytotoxic T lymphocyte responses in patients with acute hepatitis. J. Clin. Invest. 1997, 100: 503-513.

52. Bertoletti, A., Costanzo, A., Chisari, F.V., Levrero, M., Artini, M., Sette, A, *et al*. Cytotoxic T lymphocyte response to a wild type hepatitis B virus epitope in patients chronically infected by variant viruses carrying substitutions within the epitope. J. Exp. Med. 1994, 180: 933-943.

53. Shi, T.D., Wu, Y.Z., Jia, Z.C., Zhou, W., Zou, L.Y. Therapeutic polypeptides based on HBcAg (18-27) CTL epitope can induce antigen-specific CD(8)(+) CTL-mediated cytotoxicity in HLA-A2 transgenic mice. World. J. Gastroenterol. 2004, 10: 1222-1226.

54. Li, R.B., Chen, H.S., Xie, Y., Fei, R., Cong, X., Jiang, D., *et al*. Dendritic cells from chronic hepatitis B patients can induce HBV-antigen specific T cell responses. World. J. Gastroenterol. 2004, 10: 1578-1582.

55. Locarnini, S., McMillan, J., Bartholomeusz, A. The hepatitis B virus and common mutants. Semin. Liver. Dis. 2003, 23: 5-20.

56. Wei, Y., Tavis, J.E., Ganem, D. Relationship between viral DNA synthesis and virion envelopment in hepatitis B viruses. J. Virol. 1996, 70: 6455-6458.

57. Seeger, C., Mason, W.S. Replication of the hepatitis virus genome. In: DePhamphilis ML (ed), DNA replication in eukaryotic cells, 1996, p1815-1831. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

58. Sakugawa, H., Nakasone, H., Nakayoshi, T., Orito, E., Mizokami, M., Yamashiro, T., *et al*. Preponderance of hepatitis B virus genotype B contribution to a better prognosis of chronic HBV infection in Okinawa, Japan. J. Med. Virol. 2002, 67: 484-489.

59. Ganem, D., Pollack, J.R., Tavis, J. Hepatitis B virus reverse transcriptase and its many roles in hepadnaviral genomic replication. Infect. Agent. Dis. 1994, 3: 85-93.

60. Chan, H.L.Y. Significance of hepatitis B virus genotypes and mutations in the development of hepatocellular carcinoma in Asia. J. Gastroenterol. Hepatol. 2011, 26: 8-12.

61. Tagawa, M., Omata, M., Okuda, K. Appearance of viral RNA transcripts in the early stage of duck hepatitis B virus infection. Virology. 1986, 152: 477-482.

62. Chisari, F.V. Viruses, immunity, and cancer: lessons from hepatitis B. Am. J. Pathology. 2000, 156: 1117-1132.

63. Pawlotsky, J.M. The concept of hepatitis B virus mutant escape. J. Clin. Virol. 2005; 34: S125-129.

64. Okamoto, H., Tsuda, F., Sakugawa, H., Sastrosoewignjo, R.I., Imai, M., Miyakawa, Y. *et al*. Typing hepatitis B virus by homology in nucleotide sequence: comparison of surface antigen subtypes. J. Gen. Virol. 1988, 69: 2575-2583.

65. Huy, T.T.T., Ngoc, T.T., Abe, K. New complex recombinant genotype of Hepatitis B Virus identified in Vietnam. J. Virol. 2008, 82: 5657-5663.

66. Tatematsu, K., Tanaka, Y., Kurbanov, F., Sugauchi, F., Mano, S., Maeshiro, T., *et*

al. A genetic variant of Hepatitis B Virus divergent from known human and ape genotype isolated from a Japanese patient and provisionally assigned to new genotype J. J Virol. 2009, 83: 10538-10547.

67. Sakamoto, T., Tanaka, Y., Orito, E., Co, J., Clavio, Sugauchi, F., Ozasa, A., *et al*. Novel subtypes (subgenotypes) of hepatitis B virus genotypes B and C among chronic liver disease patients in the Philippines. J. Gen. Virol. 2006, 87: 1873-1882.

68. Lusida, M.I., Nugrahaputra, V.E., Soetjipto, R., Handajani, M., Nagano-Fujii, M., Sasayama, *et al*. Novel subgenotypes of hepatitis B virus genotypes C and D in Papua, Indonesia. J. Clin. Microbiol. 2008, 46: 2160-2166.

69. Nurainy, N., Muljono, D.H., Sudoyo, H., Marzuki, S. Genetic study of hepatitis B virus in Indonesia reveals a new subgenotype of genotype B in east Nusa Tenggara. Arch. Virol. 2008, 153: 1057-1065.

70. Mulyanto, Depamede, S.N., Surayah, K., Tjahyono, A.A.H., Jirintai, Nagashima, S., *et al*. Identification and characterization of novel hepatitis B virus subgenotype C10 in Nusa Tenggara, Indonesia. Arch. Virol. 2010, 155: 705-715.

71. Mulyanto, Depamede, S.N., Wahyono, A., Jirintai, Nagashima, S., Takahashi, *et al.* Analysis of the full-length genomes of novel hepatitis B virus subgenotypes C11 and C12 in Papua, Indonesia. J. Med. Virol. 2011, 83: 54-64.

72. Mulyanto, Pancawardani, P., Depamede, S.N., Wahyono, A., Jirintai, Nagashima, S., *et al*. Identification of four novel subgenotypes (C13-C16) and two inter-genotypic recombinants (C12/G and C13/B3) of hepatitis B virus in Papua Province, Indonesia. Virus. Res. 2012, 163:129-140.

73. Norder, H., Couroucé, A.M., Magnius, L.O. Complete genome, phylogenetic relatedness, and structural proteins of six strains of the hepatitis B virus, four of which represent two new genotypes. Virology. 1994, 198: 489-503.

74. Simmonds. Reconstructing the origins of human hepatitis viruses. Phil. Trans. R. Soc. Lond. B 2001, 356: 1013-1026.

75. Miyakawa, Y., Mizokami, M. Classifying Hepatitis B Virus Genotypes. Intervirology. 2003, 46: 329-338.

76. Okamoto, H., Imai, M., Tsuda, F., Tanaka, T., Miyakawa, Y., Mayumi, M. Point mutation in the S gene of hepatitis B virus for a d/y or w/r subtypic change in two blood donors carrying a surface antigen of compound subtype adyr or adwr. J. Virol. 1987, 61: 3030–3034.

77. Couroucé-Pauty, A.M., Lemaire, J.M., Roux, J.F. New hepatitis B surface antigen subtypes inside the *ad* category. Vox. Sang. 1978, 35: 304-308.

78. Norder, H., Couroucé, A.M., Magnius, L.O. Molecular basis of hepatitis B virus serotype variations within the four major subtypes. J. Gen. Virol. 1992, 73: 3141-3145.

79. Sastrosoewignjo, R.I., Sandjaya, B., Okamoto, H. Molecular epidemiology of hepatitis B virus in Indonesia. J. Gastroenterol. Hepatol. 1991, 6: 491-498.

80. Mulyanto, Tsuda, F., Karossi, A.T., Soewignjo, S., Roestamsjah, Sumarsidi, D., *et al*. Distribution of the hepatitis B surface antigen subtypes in Indonesia: implication of ethnic heterogeneity and infection control measures. Arch. Virol. 1997, 142:

2121-2129.

81. Sprengel, R., Varmus, H.E., Ganem, D. Homologous recombination between hepadnaviral genomes following in vivo DNA transfection: implications for studies of viral infectivity. Virology. 1987, 159: 454–456.

82. Owiredu, W.K., Kramvis, A., Kew, M.C. Hepatitis B virus DNA in serum of healthy black African adults positive for hepatitis B surface antibody alone: possible association with recombination between genotypes A and D. J. Med. Virol. 2001, 64: 441–454.

83. Sugauchi, F., Kumada, H., Sakugawa, H., Komatsu, M., Niitsuma, H,, Watanabe, H., *et al*. Two Subtypes of Genotype B (Ba and Bj) of Hepatitis B Virus in Japan. Clin. Infect. Dis. 2004, 38: 1222-1228.

84. Zoulim, F. Mechanism of viral persistence and resistance to nucleoside and nucleotide analogs in chronic hepatitis B virus infection. Antiviral. Res. 2004, 64: 1-15.

85. Cao, G.W. Clinical relevance and public health significance of hepatitis B virus genomic variations. W. J. Gastroenterol. 2009; 15: 5761-5769.

86. Zuckerman, A.J. Effect of hepatitis B virus mutants on efficacy of vaccination. Lancet. 2000, 355:1382–1383.

87. Ie, S.I., Thedja, M.D., Roni, M., Muljono, D.H. Prediction of conformational changes by single mutation in the hepatitis B virus surface antigen (HBsAg) identified in HBsAg-negative blood donors. Virology. 2010, 7: 326-334.

88. Carman, W.F., Zanetti, A.R., Karayiannis, P., Waters, J., Manzillo, G., Tanzi, E., *et al*. Vaccine-induced escape mutants of hepatitis B virus. Lancet. 1990, 336: 325-329.

89. Zuckerman, A.J., Harrison, T.J., Oon, C. Mutants in the S region of hepatitis B virus. Lancet. 1994, 343: 737-738.

90. Karthigesu, V.D., Allison, L.M.C., Fortuin, M., Mendy, M., Whittle, H.C., Howard, C.R. A novel hepatitis B virus variant in the sera of immunized children. J. Gen. Virol. 1994, 73: 445-448.

91. McMahon, G., Ehrlich, P.H., Moustafa, Z.A., McCarthy, L.A., Dottavio, D., Tolpin, M.D., *et al*. Genetic alterations in the gene encoding the major HBsAg: DNA and immunological analysis of recurrent HBsAg derived from monoclonal antibody-treated liver transplant patients. Hepatology. 1992, 15: 756-766.

92. Jeantet, D., Chemin, I., Mandrand, B., Tran, A., Zoulim, F., Merle, P., *et al*. Cloning and expression of surface antigens from occult chronic hepatitis B virus infections and their recognition by commercial detection assays. J. Med. Virol. 2004, 73: 508-515.

93. Gerlich, W.H., Glebe, D., Schuttler, C.G. Deficiencies in the standarization and sensitivity of diagnostic test for hepatitis B virus. J. Viral. Hepat. 2007, 14 (Suppl I): 16-21.

94. Liu, C.J., Chen, D.S., Chen, P.J. Epidemiology of HBV infection in Asian blood donors: emphasis on occult hepatitis B infection and the role of NAT. J. Clin. Virol. 2006, 36(Suppl I): S33-S44.

95. Thedja, M.D., Roni, M., Harahap, A.R., Siregar, N.C., Ie, S.I., Muljono, D.H. Occult hepatitis B in blood donors in Indonesia: altered antigenicity of the hepatitis B virus surface protein. Hepatol. Int. 2010, 4: 608-614.

96. Yuan, Q., Ou, S.H., Chen, C.R., Ge, S.X., Pei, B., Chen, Q.R., *et al*. Molecular characteristics of occult hepatitis B virus from blood donors in Southeast China. J. Clin. Microb. 2010, 48: 357-362.

97. Liu, Y., Li, P., Li, C., Zhou, J., Wu, C., Zhou, Y.H. Detection of hepatitis B virus DNA among accepted blood donors in Nanjing, China. Virology. J. 2010, 7: 193-199.

98. Hoofnagle, J.H., Seeff, L.B., Bales, Z.B., Zimmerman, H.J. Type B hepatitis after transfusion with blood containing antibody to hepatitis B core antigen. N. Engl. J. Med. 1978; 298: 1379–1383.

99. Raimondo, G., Pollicino, T., Cacciola, I., Squadrito, G. Occult hepatitis B virus infection. J Hepatology 2007; 46: 160-170.

100. Liu, S., Zhang, H., Gu, C., Yin, J., He, Y., Xie, J., *et al*. Association between hepatitis B virus mutations and the risk of hepatocellular carcinoma: a meta-analysis. J. Natl. Cancer. Inst. 2009, 101: 1066-1082.

101. Zhang, Q., Cao, G. Genotypes, mutations, and viral load of hepatitis B virus and the risk of hepatocellular carcinoma. Hepat. Mon. 2011, 11: 86-91.

102. Mun, H.S., Lee, S.A., Kim, H., Hwang, E.S., Kook, Y.H., Kim, B.J. Novel F141L Pre-S2 mutation in hepatitis B virus increases the risk of hepatocellular carcinoma in patients with chronic genotype C infections. J. Virol. 2011, 85: 123-132.

103. Omata, M. , Ehata, T. , Yokosuka, O. , Hosoda, K. , Ohto, M. Mutations in the precore region of hepatitis B virus DNA in patients with fulminant and severe hepatitis . N. Engl. J. Med. 1991; 324 (24): 1699 – 704.

104. Liang, T.J., Hasegawa, K., Rimon, N., Wands, J.R., Ben Porath, E. A hepatitis B virus mutant associated with an epidemic of fulminant hepatitis. N. Engl. J. Med. 1991; 324 (24): 1705 – 1709.

105. Carman, W.F., Jacyna, M.R., Hadziyannis, S., Karayiannis, P., McGarvey, M.J., Makris, A., *et al*. mutation preventing formation of hepatitis B e antigen in patients with chronic hepatitis B infection. Lancet. 1989, 2: 588-591.

106. Yang, H.I., Lu, S.N., Liaw, Y.F., You, S.L., Sun, C.A., Wang, L.Y., *et al*. Hepatitis B e antigen and the risk of hepatocellular carcinoma. N. Engl. J. Med. 2002; 347: 168-174.

107. Parekh, S., Zoulim, F., Ahn, S.H., Tsai, A., Li, J., Kawai, S., *et al*. Genome replication, virion secretion, and e antigen expression of naturally occurring hepatitis B virus core promoter mutants. J. Virol. 2003; 77: 6601-6612.

108. Li, D., Gu, H.X., Zhang, S.Y., Zhong, Z.H., Zhuang, M., Hattori, T. YMDD mutations and genotypes of hepatitis B virus in northern China. Jpn. J. Infect. Dis. 2006, 59: 42-45.

109. Lai, C.L., Gane, E., Liaw, Y.F., Hsu, C.W., Thongsawat, S., Wang, Y., *et al*. Telbivudine versus lamivudine in patients with chronic hepatitis B. N. Engl. J. Med. 2007; 357: 2576-2588.

110. Allen, M.I., Deslauriers, M., Andrews, C.W., Tipples, G.A., Walters, K.A., Tyrrell, D.L., *et al*. Identification and characterization of mutation in hepatitis B virus resistant to lamivudine. Lamivudine Investigation Group. Hepatology. 1998, 27: 1670-1677.

111. Toressi, J., Earnest-Silveira, L., Civitico, G., Walters, T.E., Lewin, S.R., Fyfe, J., *et al*. Restoration of replication phenotype of lamivudine-resistant hepatitis B virus mutants by a compensatory changes in the "fingers" subdomain of the viral polymerase selected as a consequence of mutations in the overlapping S gene. Virology. 2002, 299: 88-89.

112. Schildgen, O., Sirma, H., Funk, A., Olotu, C., Wend, U.C., Hartmann H, *et al*. Variant of hepatitis B virus with primary resistance to adefovir. *N* Engl. J. Med. 2006; 354: 1807-1812.

113. Angus, P., Vaughan, R., Xiong, S., Yang, H., Delaney, W., Gibbs, C., *et al*. Resistance to adefovir dipivoxil therapy associated with the selection of a novel mutation in the HBV polymerase. Gastroenterology. 2003;125: 292-297.

114. Powell, E.E., Edwards-Smith, C.J., Hay, J.L., Clouston, A.D., Crawford, D.H., Shorthouse, C., *et al*. Host genetic factors influence disease progression in chronic hepatitis C. Hepatology. 2000; 31: 828-833.

115. Wang, F.S. Current status and prospects of studies on human genetic alleles associated with hepatitis B virus infection. World. J. Gastroenterol. 2003; 9 (4): 641-644.

116. Thursz, M.R., Kwiatkowski, D., Allsopp, C.E.M., Greenwood, B.M., Thomas, H.C., Hill, A.V.S. Association between an MHC class II allele and clearance of hepatitis B virus in the Gambia. N. Engl. J. Med. 1995; 332(16): 1065–1069.

117. Ahn, S.H., Han, K.H., Park, J.Y., Lee, C.K., Kang, S.W., Chon, C.Y., *et al*. Association between hepatitis B virus infection and HLA-DR type in Korea. Hepatology. 2000; 31: 1371-1373.

118. Thio, C.L., Thomas, D.L., Karack, P., et al. Comprehensive analysis of class I and class II HLA antigens and chronic hepatitis B virus infection. J Virol 2003;77(22):12083–12087.

119. Almarri, A., Batchelor, J.R. HLA and hepatitis B infection. Lancet. 1994; 344(8931): 1194–1195.

120. Thio, C.L., Carrington, M., Marti, D., O'Brien, S.J., Vlahov, D., Nelson, K.E., *et al*. Class II of HLA alleles and hepatitis B Virus persistence African Americans. J. Infect. Dis. 1999; 179: 1004-1006.

121. Jiang, Y.G., Wang, Y.M., Liu, T.H., Liu, J. Association between HLA class II gene and susceptibility or resistance to chronic hepatitis B. World. J. Gastroenterol. 2003; 9(10): 2221–2225.

CHAPTER 2

Ethnogeographical structure of hepatitis B virus genotype distribution in Indonesia and discovery of a new subgenotype, B9

Meta D. Thedja[1,3], David H. Muljono[1], Neni Nurainy[1,2,] Caecilia H.C. Sukowati[1], Jan Verhoef[3] and Sangkot Marzuki[1,4]

[1]Eijkman Institute for Molecular Biology, Jl. Diponegoro 69, Jakarta 10430, Indonesia

[2]PT Bio Farma, Jl. Pasteur 28, Bandung 40161, Indonesia

[3]Eijkman Winkler Institute, Utrecht Medical Centre, Utrecht, The Netherlands

[4]Department of Medicine, Monash Medical Centre, Monash University, Clayton, Vic 3168, Australia

Arch Virol (2011); 156(5):855-68

Abstract

The distribution of Hepatitis B Virus (HBV) in the populations of island Southeast Asia is of medical and anthropological interests, and is associated with an unusually high genetic diversity. This study examined the association of this HBV genetic diversity with the ethnogeography of the populations of the Indonesian archipelago. Whole genome analysis of 21 HBV isolates from East Nusa Tenggara and Papua revealed two recently reported HBV/B subgenotypes unique to the former, B7 (7) and B8 (5), and uncovered a further novel subgenotype designated B9 (4). Further isolates were collected from 419 individuals with defined ethnic backgrounds representing 40 populations. HBV/B was predominant in Austronesian languages speaking populations, whereas HBV/C was major in Papua and Papua influenced populations of Moluccas; HBV/B3 was the predominant subgenotype in the western half of the archipelago [speakers of the Western Malayo-Polynesian (WMP) branch of Austronesian languages], whereas B7, B8 and B9 were specific for Nusa Tenggara [Central Malayo-Polynesian (CMP)]. The result provides the first direct evidence that the distribution of HBV genotypes/subgenotypes in the Indonesian archipelago is related with the ethnic origin of its populations, and suggests that the HBV distribution is associated with the ancient migratory events in the peopling of the archipelago.

Keywords: HBV genotype/subgenotype, distribution, ethnogeographical, Indonesia, human migration.

Introduction

Hepatitis B virus (HBV) is a major cause of liver diseases, particularly in Asia. Genetic variability of HBV plays an important role in the development to chronic hepatitis B, and is associated with the clinical outcome and response to treatment [17, 38, 56]. Eight HBV genotypes A to H have been identified [2, 24, 25, 27, 30, 31, 36, 45], with genotype B and C predominant among Asian populations. Very recently, two new additional HBV genotypes, HBV/I and HBV/J, were proposed for isolates collected from Laos and Japan, respectively [14, 51].

Eight subgenotypes have been reported for the Asian HBV genotype B (HBV/B), each with different geographical predominance: B1 in Japan, B2 in China, B3 in Indonesia, B4 in Vietnam, B5 in the Philippines, B6 in the Arctic indigenous population, B7 and B8 in eastern Nusa Tenggara islands of Indonesia [28, 29, 32, 34, 41, 42]. Similarly, HBV genotype C (HBV/C) has been classified into six geographically related subgenotypes: C1, C5 and

C6 in Southeast Asia, and C2 in east Asia [13, 24, 29, 41, 55], C3 mostly in the Pacific, and C4 in the Aborigines of Northeast Australia [32, 46].

The distribution of HBV genotypes and subgenotypes in the populations of island Southeast Asia is of particular interests. The Indonesian part of the archipelago alone consists of approximately 17,500 islands, and is home to 230 millions people of more than 500 ethnic populations, inhabiting around 6,000 islands [48]. The main origins of these populations are believed to be two major waves of ancient migration: the initial peopling of the archipelago by modern human 60,000 years before present (yBP) and the arrival of Austronesian languages speakers around 5,000 yBP [3]. Information regarding the distribution of HBV genotypes/subgenotypes amongst the ethnic populations of the archipelago, therefore, might reveal knowledge of anthropological significance. Such information is of medical importance, as this ethnically diverse region is now the major source of migrant populations in the more developed countries.

Our recent study suggests that the HBV genotype/subgenotype distribution in this archipelago is complex, and indeed associated with the ethnic background of the populations rather than with geographical locations [34]. For example, HBV/B3 is found mainly in ethnic populations of the western half of the archipelago while HBV/B7 is associated with ethnic populations of the Nusa Tenggara islands of the eastern half. A recent nationwide study of HBV molecular epidemiology in Indonesia, showing geographical specificity of HBV genotypes/subgenotypes distribution, also indicated possible association with the ethnological origins of the populations [28]. This study was aimed to provide evidence that the HBV genotypes/subgenotypes distribution is indeed related to the ethnogeographical structure of the Indonesian populations, in a study involving a large number of subjects with carefully defined ethnic backgrounds representing 40 ethnic populations. Our results demonstrate the association of HBV genotypes/subgenotypes with the ethnological origins of the populations of the Indonesian archipelago.

Material and methods

Serum samples and ethnic populations

A total of 440 serum samples positive for HBsAg (310 men and 130 women; mean age, 40.2 $\pm$ 5.2 years) were obtained from asymptomatic carriers (263 samples), HBV-related liver disease patients (158 samples) who never received antiviral therapy, and blood donors (19 samples). The samples were collected from 20 geographical locations (Table 1). None of the participants was coinfected with either hepatitis C virus or human

immunodeficiency virus. The ethnic background of the individuals from whom the samples were obtained was carefully documented and ascertained for at least three previous generations, both maternally and paternally as previously described [26].

Ethnic populations were selected to represent the clustering of their genetic and linguistic affinities based on the mapping of human genetic diversity in Asia by the HUGO Pan-Asian SNP Consortium [12, 33, 52]: the Austronesian languages-speaking populations of western islands of Indonesia (Sumatra, Kalimantan and Java), the Austronesian-speaking populations of the islands of Sulawesi and Nusa Tenggara archipelago, and the Papua and Papuan-speaking eastern island populations. The origins and characteristics of the individuals from whom the HBV isolates obtained are shown in Table 1. Samples from Indonesians of Chinese ethnic origin were collected in three big cities (Jakarta, Surabaya and Medan). The study was approved by the Eijkman Institute Research Ethics Commission (EIREC No. 23/2007).

HBV genome sequencing

Viral DNA was extracted from 140 µL of HBsAg positive serum using QIAamp® DNA Mini Kit (Qiagen Inc., Chatsworth, CA) according to the manufacturer's instruction. HBV DNA was detected by nested PCR using Platinum®*Tag* DNA Polymerase (Invitrogen), targeting the conserved segment within the S gene, using primer sets as described previously [36, 37, 58]: S2-1 (5'-CAAGGTATGTTGCCCGTTTG-3', nt 455-474) and S1-2 (5'-CGAACCACTGAACAAATGGC-3', nt 704-685) for the first round, and S088 (5'-TGTTGCCCGTTTGTCCTCTA-3', nt 462-471) and S2-2 (5'-GGCACTAGTAAACTGAGCCA-3', nt 687-668) for the second round. Denaturizing, annealing and extension were carried out at 94°C for 30 s, 55°C for 30 s and 72°C for 1 min, respectively for both rounds of PCR (35 cycles for the first and 25 for the second rounds).

For whole genome sequencing, five overlapping fragments were first amplified using primer sets described previously [35, 47, 49]: PS8-1 (5'-GTCACCATATTCTTGGGAAC-3') and HS6-2 (5'-GCCAAGTGTTTGCTGACGCA-3') for fragment A (nt 2,817 – 1,194), S2-1 and HB4R (5'-CGGGACGTAGACAAAGGACGT-3') for fragment B (nt 487 – 1,434), HB5F (5'-GCATGGAGACCACCGTGAAC-3') and S013 (5'-TCCACAGAAGCTCCAAATTCTTTT-3') for fragment C (nt 1,256 – 1,941), PC1 (5'-CATAAGAGGACTCTTGGACT-3') and HB9R (5'-GGATAGAACCTAGCAGGCAT-3') for fragment D (nt 1,653 – 2,656), and HB10F (5'-CGCAGAAGATCTCAATCTCGG-3') and T734 (5'-CTTCCTGACTGSCGATTGG-3') for fragment E (nt 2,417 – 3,156). The amplification reaction was carried out for 35 cycles of denaturation at 94°C for 30 s,

annealing at 55-59°C (depending of the primer pair used) for 30 s, and extension at 72°C for 60 s, and elongation at 72°C for 7 min. Amplification products were sequenced directly using the Big Dye Terminator Reaction kits on an ABI 3130 genetic analyzer (ABI Perkin Elmer, Norwalk, CT, USA). The genome sequences were assembled and analyzed using BioEdit version 7.0.5 software.

For the sequencing of the Pre-S2 region, semi-nested PCR was carried out employing primer sets PS1-1 (5'-CCTCCTGCCTCCACCAATCG-3', nt 3125-3144) and t703 (5'-CAGAGTCTAGACTCGTGGTG-3', nt 242-261) for the first round, and PS1-1 and PS5-2 (5'-CTCGTGTTACAGGCGGGGTT-3', nt 190-210) for the second round [49]. The PCR was carried out for 35 cycles of denaturation at 94°C for 30 s, annealing at 57-59°C (depending of the primer pair used) for 30 s, and extension at 72°C for 60 s, and elongation at 72°C for 7 min.

HBV genotype and subgenotype determination

Twenty-four complete genome sequences generated in this study, together with 141 obtained from GenBank (including Indonesian isolates recently reported [28, 55] were aligned. Phylogenetic tree was constructed and genetic distance was calculated with the 6-parameter method [40]. The genotypes and subgenotypes of the 24 new isolates were determined based on their phylogenetic co-clustering with the previously defined sequences. For the wider study of the HBV ethnogeographical distribution, genotype and subgenotype assignment were carried out for 654 isolates based on their preS2 sequences (440 sequences generated in this study and 214 sequences from GenBank) employing signatures of specific Single Nucleotide Polymorphisms (SNPs) diagnostic for the various Asian HBV genotypes and subgenotypes, as previously reported [34] and further developed in this study (Table 2).

Results

HBV genotypes and the discovery of another novel of HBV/B subgenotype from Nusa Tenggara islands

Phylogenetic analysis of the 24 complete HBV genomes obtained in this study [21 sequences from ethnic populations of the eastern region: Sumbanese (7), Flores (8), Alorese (1), and Papuan (Merauke 3, Jayapura 1 and Sentani 1; and 3 sequences from ethnic populations of the western region: Javanesse (1) and Minang (2)], along with 141 sequences from GenBank, identified 3 HBV genotypes, 17 HBV/B, 6 HBV/C, and 1

HBV/D (Figure 1). Of the 17 HBV/B isolates, 1 belonged to B3, 7 to B7, 5 to B8, but 4 to an unclassified cluster. The latter was distinct from the existing HBV/B1-B8 with significant posterior probability (100). Phylogenetic trees constructed from ORFs P and S were consistent with that obtained from the complete genome, although discordant with that constructed from ORF C as previously observed, as the consequence of recombination event involving this region [47]. Together with an intersubgenotype divergence of more than 4% to each of B1, B2, B4 and B6 (Table 3), we propose that the unidentified cluster represents a novel subgenotype, designated B9 (Figure 1).

This B9 was distinguished from other HBV/B subgenotypes by specific features seen in the region encoding HBsAg and HBcAg. In the part of S gene (nt 155 – 832) encoding small surface protein (226 amino acid residues), HBsAg, two nucleotide substitutions were found in the B9 subgenotype isolate group, which are not present in other isolate groups of B subgenotypes (Supplementary Figure 1). These nucleotide substitutions were 555A and 570T, both substitutions caused silent mutation. Within their core regions, encompassing nt 1901 - 2452, six nucleotide substitutions which are not present in other isolate groups of HBV/B subgenotypes were identified, 49G, 207A, 214T, 228G, 229A, and 291A (Supplementary Figure 2). Three amino acid substitutions, which are not present in other isolate groups of HBV/B subgenotypes, were found, Val^{15}, Leu^{72}, and Lys^{77}.

The phylogenetic relationship of Asian HBV genotypes/subgenotypes with ethnic origin and geographical distribution was also demonstrated that HBV/B1 and B2, which is found in Japan and China, respectively, clearly separated with the other HBV/B from island southeast Asia (B3, B4, B5, B7, B8 and B9). Further interesting, HBV/C1, C2 and C5, predominant in southeast and east Asia, clustered in one major cluster completely distinct from C6, specific for Papua. The other HBV/C subgenotypes, specific for the Oceanian (C3) and Aboriginal populations of northern Australia (C4) formed distant clusters.

Ethnogeographical distribution of HBV/B, HBV/C and HBV/D subgenotypes in the Indonesian archipelago

HBV genotypes/subgenotypes in this study were determined based on the diagnostic sites of SNPs of the PreS2 sequence. The genotypes of 440 HBV isolates were 312 HBV/B (70.9%), 121 HBV/C (27.5%), and 7 (1.6%) HBV/D (Table 4). The distribution of the HBV genotypes and their subgenotypes showed distinct ethnic-related patterns of the prevalence of genotypes B, C and D (Figure 2).

Of the 189 isolates from the islands of western Indonesia (Sumatra, Nias, Mentawai, Kalimantan, Java and Lombok islands), HBV/B accounted for almost 74.6% (141 isolates) followed by HBV/C (48 isolates; 25.4%). At the subgenotype level, B3 was by far predominant (70.9%), followed by B8 (9.9%), B9 (7.8%), and B5 (6.4%), while B2 and B7 represented only 2.8% and 2.1% of the total HBV/B, respectively. Of the HBV/C, C1 was detected in 28 (58.3%) and C2 in 20 (41.7%) isolates. Thus, HBV/B and its subgenotype B3 were the predominant genotype and subgenotype in the islands of western Indonesia, with the exception of the Minang population of west Sumatra, in which HBV/C and its subgenotype C1 were the predominant genotype and subgenotype.

In contrast, in the Moluccas and Papua, in the far east of the Indonesian archipelago, HBV/C was the predominant genotype (80%), followed by HBV/D (16.7%), with noticeably only one HBV/B detected out of the 30 isolates examined. C1 was found in 37.5%, C2 in 20.8%, C5 in 12.5%, and C6 in 29.2% of the 24 HBV/C isolates. In the coastal populations of Papua, HBV/C6, a recently reported HBV/C subgenotype [24, 28, 55], was by far the major subgenotype (43.8%) with C2 being the second (31.3%), while HBV/D constituting 25%.

In between, in Sulawesi and East Nusa Tenggara (Sumba, Flores and Alor) islands, all the three HBV genotypes, B, C and D, were detected in 147 (78.2%), 39 (20.7%), and 2 (1.1%) of the 188 isolates examined, respectively. More variation in HBV/B subgenotypes were detected: B3 in 12 (8.2%), B5 in 15 (10.2%), B7 in 64 (43.5%), B8 in 11 (7.5%) and B9 in 45 (30.6%). Of HBV/C, C1, C2 and C5 accounted for 21 (53.9%), 13 (33.3%) and 5 (12.8%) isolates, respectively.

The distribution of HBV genotypes/subgenotypes among Indonesian of Chinese ethnic origin was dominated by HBV/B2 (42.4% of the total 33 HBV isolates), followed by HBV/B3 as the second dominant HBV/B subgenotype (25.5%).

The 24 HBV complete genomes together with 416 PreS2 sequences obtained in this study have been deposited in the GenBank with accession numbers GQ358136 to GQ358159 and GU071282 to GU071721, respectively.

Distribution of hepatitis B virus genotypes in Indonesia

Table 1. HBV isolates collected in the present study

Geographical location	Ethnic background	n	Gender (M/F)	Mean age	Accession Number
I. West Indonesia (Austronesian Cluster 1)		**172**	**102/70**	**44.3 ± 11.6**	
North Sumatra	Karo Batak	14			GU071439-GU071452
West Sumatra	Minang	51			GU071537-GU071587
South Sumatra	Malay	54			GU071460-GU071513
Nias	Nias	8			GU071588-GU071595
Mentawai	Mentawai	6			GU071527-GU071532
Central Java	Javanese	35			GU071398-GU071432
East Kalimantan	Dayak Benuaq	4			GU071307-GU071310
II. East Indonesia – Sulawesi & Nusa Tenggara (Austronesian Cluster 2)		**188**	**142/46**	**45.1 ± 14.9**	
Lombok	Sasak	17			GU071612-GU071628
North Sulawesi	Minahasa	4			GU071533-GU071536
South Sulawesi	Mandar	13			GU071514-GU071526
	Torajan	5			GU071635-GU071639
	Kajang	6			GU071433-GU071438
	Makassar	7			GU071453-GU071459
West Sumba	Kodi	12			GU071678-GU071689
	Lamboya	18			GU071690-GU071707
	Loli	4			GU071708-GU071711
	Anakalang	8			GU071664-GU071671
	Mamboro	6			GU071712-GU071717
	Wanokaka	4			GU071662-GU071663, GU071720-GU071721
	Mbilur/Pangadu	2			GU071676-GU071677
	Waimangura	2			GU071718-GU071719
	Bukambero	4			GU071672-GU071675
East Sumba	Kambera	9			GU071356-GU071364
West Flores	Larantuka	4			GU071640, GU071642-GU071644
	Lembata	1			GU071645
	Pantar	17			GU071641, GU071646-GU071661

Table 1. continued

Geographical location	Ethnic background	n	Gender (M/F)	Mean age	Accession Number
East Flores	LIO Selatan	9			GU071336-GU071344
	LIO Tengah	3			GU071345-GU071347
	Bere	2			GU071311, GU071313
	Rampasasa	1			GU071348
	Boawae	2			GU071314-GU071315
	Soa	5			GU071349-GU071353
	Wogo	3			GU071312, GU071354-GU071355
	Cibal	4			GU071316-GU071319
	Flores Timur	16			GU071320-GU071335
III. East Indonesia - Papua & Moluccas (Papuan speaking and influenced Cluster)		**47**	**47/0**	**36.5 ± 13.9**	
Alor	Alorese	17			GU071282-GU071298
Papua	Papuan	16			GU071596-GU071611
Moluccas	Ternate	6			GU071629-GU071634
	Ambonesse	8			GU071299-GU071306
IV. Jakarta, Surabaya and Medan (Indonesian Chinese)	Han Chinese	**33**	**19/14**	**35 ± 12.4**	GU071365-GU071397

The HBV isolates were collected of 440 individuals from 40 different ethnic populations of the Indonesian archipelago. The ethnic background of the individuals from whom the samples were obtained was carefully documented and ascertained for at least three previous generations, both maternally and paternally as previously described [26]. The isolates were arranged into four major groups based on the genetic clustering [52] of the ethnic backgrounds of the individuals from which they were isolated, which has been shown [52] to be is consistent with their linguistic clustering [54].

Table 2. HBV genotype and subgenotype determination based on diagnostic SNPs signatures in the PreS2 Region

a. HBV Genotype Determination

	Specific SNPs								
	20	25	27	43	45	76	96	135	150
Genotypes B	A	T	A	G	C	A	A	C	C
C	G	G	T	G	G	C	C	C	C
D	A	G	A	A	G	A	C	T	T

b. HBV/B Subgenotyping

	Specific SNPs																
	10	**31**	**34**	35	46	55	85	87	93	99	100	105	**109**	110	**123**	128	148
Subgenoypes B1	T	T	C	A	T	T	A	G	C	A	T	T	A	G	T	T	G
B2		T	C	A	C	T	G	G	C	A	T	T	T	G	T	T	A
B3		T	C	A	T	T	A	A	T	A	C	C	T	T	T	T	G
B4		T	C	A	T	T	G/A	G/A	C	A	T	C	T	G	C	T	A
B5		T	C	A	T	T	G	A	C	A	C	C	T	T	T	T	A
B6		T	C	A	T	T	G	A	C	T	T	C	T	T	T	C	A
B7		T	C	A/**C**	T	C/**T**	G	A	T	C	C	C	T	T	T	T	G
B8		T	G	A	T	T	G	A	T	A	C	C	T	T	C	T	A
B9		C	C	A	T	T	A	A	T	A	C	C	T	T	T	T	A

d. HBV/D Subgenotyping

	Specific SNPs			
	37	105	111	113
Subgenoypes D1	A	T	A	A
D2	G	C	C	A
D3	A	C	C	T
D4	A	C	C	A

Table 2. continued

c. HBV/C Subgenotyping

	Specific SNPs																			
	6	7	8	13	16	27	28	40	46	**49**	84	99	105	**109**	111	115	**132**	**147**	149	154
Subgenotypes C1	G	C	A	T	C	T	A	G	C	A	T	C	C	C	C	A	**G**	**C**	C	T
C2	C	A/**C**	A	T	C	T	A	G	C	A	T	C	C	A	C	A	**G**	**C**	C	C
C3	C	A	A	T	A	T/A	A/G	G/C	C	A/T	T	C	C	C	T	T	G	T	A	C
C4	C	A	G	T	C	A	G	A	T	G	T	A	T	C	A	C	A	C	T	C
C5	C	A	A	C	T	T	A	A	C	A	C	C	C	C	C	A	G	C	C	C
C6	C	A	G	C/T	C	T	A	G	C	T	T	C	C	C	C	C	G	C	C	C

For genotype assignment, the sequence of the PreS2 region of the HBV genome was obtained for each isolates. The genotype was determined from the PreS2 sequence employing the nine SNPs (Table 2a) that form the signatures of HBV genotype B (20A, 25T, 27A, 43G, 45C, 76A, 96A, 135C, and 150C), C (20G, 25G, 27A, 43G, 45C, 76A, 96A, 135C, and 150C) and D (20A, 25G, 27A, 43A, 45G, 76A, 96C, 135T, and 150T) as previously reported [30]. Isolates of HBV/B, HBV/C and HBV/D genotypes were further subgenotyped based on the SNP signatures of the PreS2 region that define these subgenotypes (Table 2b, 2c and 2d respectively). Nucleotide numbering is based on the *EcoRI* endonuclease restriction site. In bolds are diagnostic sites defined in this study, additional to those reported previously [not bold; [30]].

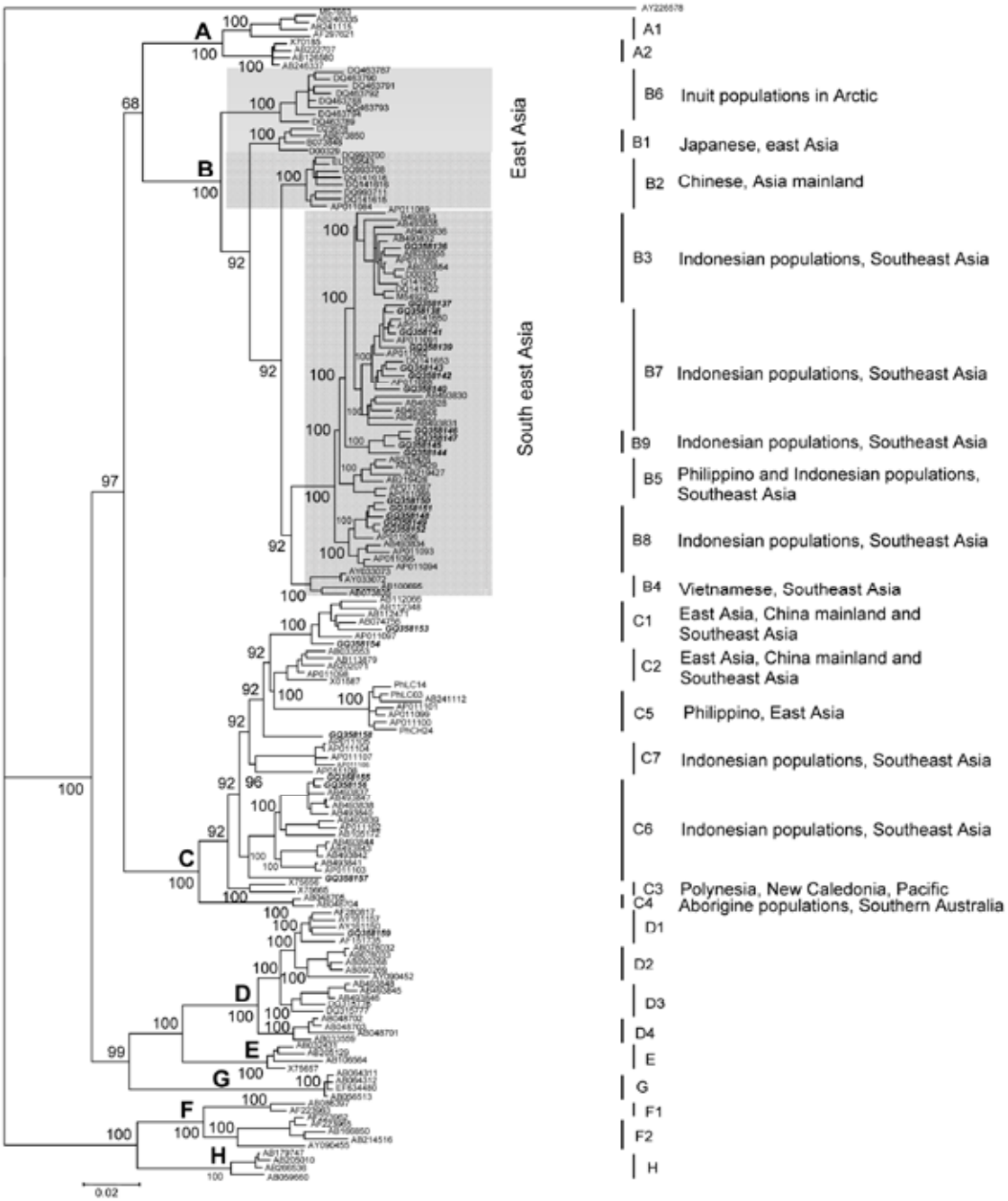


Figure 1. Phylogenetic analysis of 24 new HBV whole genome sequences revealed a novel HBV/B subgenotype. A phylogenetic tree was constructed from 24 whole genome sequences generated in the present study (***bold and italic***), and 141 representative sequences retrieved from the GenBank [8 HBV/A (A1 4 , A2 4), 58 HBV/B (B1 4, B2 8, B3 12, B4 4, B5 6, B6 8, B7 11, and B8

5), 38 HBV/C (C1 5, C2 5, C3 2 , C4 2, C5 7, C6 12, C7 5), 18 HBV/D (D1 4, D2 5, D3 5 and D4 4), 4 HBV/E, 7 HBV/F (F1 2, F2 5), 4 HBV/G and 4 HBV/H]. Genetic distance was calculated by six-parameter method (40), with the Wooly Monkey Strain (WMHBV) AY2266578 as outgroup. The length of the horizontal bar indicates the number of nucleotide substitutions per site, and the posterior probability values are indicated at the roots of the tree. The tree demonstrates the clear distinction between the east Asia and southeast Asia HBV/B subgenotype groups (the non recombinant type (42) is indicated by light shadow, while the recombinant type by the darker shadow), and revealed that 17 out of the 24 new sequences were of HBV/B (1 B3, 7 B7, 5 B8, and 4 belonging to a previously unidentified but distinct cluster), 5 HBV/C (2 C1 and 3 C6), and 1 HBV/D (D1). The unidentified cluster was separated from other HBV/B southeast Asia type with good value of posterior probability, suggesting that it is of a novel HBV/B subgenotype designated B9. Of the sequences retrieved from the GenBank, 5 reported previously as B3 (AB493827, AB493828, AB493829, AB493830, and AB493831) (55) clustered with B7. Further, one previously unidentified isolate (AB493834) (55) was found to cluster with B8.

Discussion

Several new HBV/B subgenotypes have been discovered in recent years from studies in ethnic populations of Asia, in addition to the initial subgenotypes identified in Japanese (Bj/B1), Chinese (Ba/B2) and 'Indonesian' (B3) [32, 47]. Recent studies [28, 34] suggested that the eastern islands of the Indonesian archipelago have an unusually high HBV/B genetic diversity. For example, two subgenotypes have been reported (B7 and B8) from this region, in addition to B3 that is dominant in the western half of the archipelago. This observation is in contrast with that of the Japanese and Chinese populations in east Asia, which exhibit only one HBV/B subgenotype for each population, B1 and B2, respectively [32, 42].

One more HBV/B subgenotype, B9, was discovered in the present study, in the east Nusa Tenggara islands of Indonesia. In addition to the posterior probability value (100), pairwise comparison of the HBV/B genome sequences (Table 3) revealed that the intersubgenotypic divergences of HBV/B9 against B1, B6, B2 and B4 were significantly higher than the suggested 4% as the distinguishing divergence for subgenotypes [20] (6.07 ± 0.49, 5.87 ± 0.29, 4.86 ± 0.37, and 4.82 ± 0.45, respectively). Although the divergences were less against B5, B8, B7 and B3 (3.43 ± 0.25, 3.22 ± 0.24, 3.21 ± 0.31, and 3.07 ± 0.24, respectively), we argue, that consideration of the distinct geographical and host ethnicity association [21], in addition to the phylogenetic and genetic distance data, define the unidentified cluster as a distinct subgenotype. Consistent with the above arguments, B5 initially discovered in the Philippines [41] also shows only 3.2% nucleotide divergence from the better established B3 of western Indonesia (Table 3).

This subgenotype B9 branched at a position more distant than the ancestral point of B3 and B7 (Figure 1), suggesting that it is evolutionarily older than B3 and B7.

Amino acid substitution patterns in core proteins of B9 isolates (Val[15], Leu[72], Lys[77]) also distinguished them from other HBV/B subgenotypes. Two of the three substitutions occurred at known immune recognition sites: the immunodominant CD4 T-cell epitopes (amino acids 1–20) and the B-cell determinant (amino acid 74–89) [15, 21]. Further bioinformatics and experimental studies of B9 together with other HBV/B subgenotypes would be needed to understand the dynamic interactions between the virus and host immune system as well as the natural selection in different host populations

A closer examination of the phylogenetic relationships of the 165 complete genome sequences (24 sequences from this study and 141 from the GenBank) between the various Asian HBV genotypes/subgenotypes clearly revealed their ethnogeographical association (Figure 1). HBV/B subgenotypes specific for east Asia (B1 in Japan and B2 in China) clearly separated from those of island southeast Asia (B3, B4, B5, B7, B8 and B9). This observation revealed that HBV/B1 and B2 were the HBV/B subgenotypes specific for east Asia, while B3, B4, B5, B7, B8, and B9 specific for southeast Asia.

To genotype and subgenotype large number of HBV isolates in this study (440 isolates), we have utilized the sequence diversity of the PreS2 sequence. We have shown previously that the sequence of the PreS2 region—which is more variable than the S region, perhaps because it is subject to less functional constraints —can be used for reliable HBV/B and HBV/C subgenotyping on the basis of a set of diagnostic SNPs [34]. These diagnostic SNPs were determined from PreS2 sequences of HBV isolates subgenotyped by phylogenetic analysis of whole genome sequences [34]. Additional diagnostic sites were identified from the 24 new whole genome sequences (Table 2). Thus, this study confirmed the usefulness of diagnostic SNPs of PreS2 sequence particularly for large sample analysis.

The result of our study provides the first direct evidence that the distribution of HBV genotypes/subgenotypes in the Indonesian archipelago is related to the ethnic origins of its populations. The genetic clustering of the ethnic populations of Indonesia has been defined as part of a recent large study on the genetic diversity of Asia by The HUGO Pan-Asian SNP Consortium [52]. The clustering is consistent with the ethnolinguistic structure of the Asian populations investigated [54]. Except for Papua, most of the ethnic populations of Indonesia are speakers of languages belonging to the Austronesian linguistic family. Our finding that HBV/B is the predominant genotype in the Indonesian archipelago, except in Papua and Papuan influenced neighboring populations of Moluccas where HBV/C was predominant, suggests that HBV/B is specifically

associated with the Austronesian speakers, whereas HBV/C is the major genotype in Papua.

Of particular significance in relation to the origin of the ethnic populations is the association between the observed HBV/B subgenotypes and the linguistic subgroups of the Austronesian speakers. There are three Austronesian language subgroups in the Southeast Asian archipelago [4, 54]: Western Malayo-Polynesian (WMP; Sumatra, Java, Kalimantan, Sulawesi and western islands of Nusa Tenggara), Central Malayo-Polynesian (CMP; eastern islands of Nusa Tenggara islands and south Moluccas) and South Halmahera West New Guinea (SHWNG). HBV/B3 is the major subgenotype in the Austronesian WMP speakers of the western half of Indonesia, whereas unique HBV/B subgenotypes—B7, B8, and B9—were observed in the populations of East Nusa Tenggara islands belonging to the Austronesian CMP linguistic subgroup.

The observation of HBV/B subgenotypes that are unique to the Indonesian ethnic populations and their distribution following the ethnolinguistic structure of the populations, suggests that it is unlikely that these subgenotypes have been introduced in recent times. Rather, the result indicates that the origin of HBV distribution is associated with the ancient migratory events involved in the peopling of the archipelago. Archaeological and anthropological findings indicate that there were two major migratory events associated with the peopling of the Indonesian archipelago: the first occurred some 60,000 years (yBP) with the earliest arrival of modern humans in their continuing migration from Africa to Papua and Australia; while the second occurred around 3000-5,000 yBP as part of the diaspora of Austronesian languages speaking populations [3, 11]. The Austronesian speakers replaced and perhaps assimilating most of the original Austromelanoid populations, but in the island of Papua New Guinea the populations originating from the initial peopling event some 50,000 years earlier remain isolated, separated by extreme geographical features. The long isolation is reflected by the fact that there are more than 1,000 distinct languages spoken in the island belonging to three language families [6, 7], in addition to the Austronesian languages spoken by the coastal populations.

It has been suggested that the HBV evolution history in primates is a relatively recent evident, with the divergence in humans and apes has occurred only in the last 6,000 to 7,000 years [8, 61]. However, the above suggestions are incompatible with the finding of HBV in isolated Papua New Guinea and Australia aboriginal populations [10, 43, 46]. The ubiquitous distribution of HBV/C in east and Southeast Asia, Papua New Guinea and Australia all argues for an early introduction HBV along with the initial peopling of the islands.

Table 3. Inter- and intra-subgenotypic divergence (%) of the nine HBV/B subgenotypes from 88 isolates and their country origins.

Subgenotype	B1 (9)	B2 (15)	B3 (12)	B4 (4)	B5 (7)	B6 (8)	B7 (19)	B8 (10)	B9 (4)
B1 (Japan)	**2.1 ± 0.63**	4.32 ± 0.54	5.81 ± 0.51	4.95 ± 0.65	5.96 ± 0.49	5.56 ± 0.47	6.09 ± 0.54	5.72 ± 0.55	6.07 ± 0.49
B2 (China)		**1.68 ± 0.45**	4.39 ± 0.37	3.74 ± 0.39	4.63 ± 0.34	6.19 ± 0.36	4.68 ± 0.38	4.44 ± 0.38	4.86 ± 0.37
B3 (Indonesia)			**1.32 ± 0.43**	4.61 ± 0.49	3.16 ± 0.27	5.28 ± 0.29	2.29 ± 0.33	2.75 ± 0.27	3.07 ± 0.24
B4 (Vietnam)				**2.36 ± 1.19**	4.62 ± 0.51	6.65 ± 0.49	4.75 ± 0.54	4.51 ± 0.58	4.82 ± 0.45
B5 (Philippines)					**1.84 ± 0.48**	5.69 ± 0.33	3.17 ± 0.35	2.94 ± 0.34	3.43 ± 0.25
B6 (Artic)						**1.84 ± 0.51**	5.34 ± 0.36	5.25 ± 0.31	5.87 ± 0.29
B7 (Indonesia)							**1.80 ± 0.66**	2.96 ± 0.36	3.21 ± 0.31
B8 (Indonesia)								**1.60 ± 0.55**	3.22 ± 0.24
B9 (Indonesia									**1.62 ± 0.69**

Numbers in brackets show the total isolates for each subgenotype. Intrasubgenotypic divergence is shown in bold.

Table 4. Genotype and subgenotypes distribution of 440 HBV isolates from various geographical origins in Indonesia

Genotype	Accession Number	Geographic origin / reported from
B2	GU071382, GU071383, GU071384, GU071387, GU071365, GU071367, GU071368, GU071369, GU071370, GU071372, GU071374, GU071376, GU071377, GU071380, GU071614, GU071309, GU071439, GU071573	Java, Lombok, North and West Sumatra
B3	GU071379, GU071378, GU071375, GU071371, GU071373, GU071389, GU071389, GU071390, GU071391, GU071388, GU071283, GU071287, GU071285, GU071313, GU071347, GU071626, GU071627, GU071628, GU071623, GU071612, GU071615, GU071616, GU071617, GU071618, GU071613, GU071516, GU071517, GU071520, GU071522, GU071515, GU071433, GU071459, GU071529, GU071460, GU071462, GU071463, GU071464, GU071465, GU071466, GU071467, GU071468, GU071469, GU071471, GU071472, GU071473, GU071474, GU071476, GU071477, GU071478, GU071479, GU071480, GU071481, GU071482, GU071483, GU071484, GU071485, GU071486, GU071487, GU071488, GU071489, GU071490, GU071491, GU071492, GU071493, GU071494, GU071495, GU071496, GU071497, GU071498, GU071499, GU071502, GU071504, GU071505, GU071507, GU071508, GU071509, GU071510, GU071512, GU071513, GU071440, GU071441, GU071442, GU071443, GU071444, GU071447, GU071448, GU071451, GU071571, GU071575, GU071577, GU071546, GU071555, GU071563, GU071405, GU071406, GU071407, GU071408, GU071409, GU071410, GU071411, GU071412, GU071413, GU071414, GU071415, GU071416, GU071417, GU071418, GU071420, GU071421, GU071422, GU071423, GU071424, GU071425, GU071426, GU071427, GU071428, GU071429, GU071430, GU071431, GU071398, GU071401, GU071399	Java, North and West Sumatra, Mentawai, South Sulawesi, Lombok, East Flores, Alor
B5	GU071402, GU071403, GU071404, GU071548, GU071308, GU071590, GU071592, GU071593, GU071531, GU071453, GU071455, GU071458, GU071454, GU071434, GU071436, GU071523, GU071521, GU071671, GU071677, GU071357, GU071349, GU071350, GU071352, GU071322	Java, West Sumatra, South Sulawesi, Nias, Mentawai, West and East Sumba, East Flores
B7	GU071400, GU071450, GU071310, GU071456, GU071437, GU071438, GU071514, GU071519, GU071526, GU071638, GU071636, GU071674, GU071711, GU071664, GU071670, GU071676, GU071720, GU071721, GU071662, GU071690, GU071696, GU071703, GU071704, GU071695, GU071695, GU071668, GU071718, GU071719, GU071356, GU071358, GU071363, GU071643, GU071644, GU071651, GU071336, GU071337, GU071338, GU071339, GU071340, GU071341, GU071342, GU071343, GU071344, GU071345, GU071346, GU071320, GU071321, GU071324, GU071326, GU071331, GU071334, GU071311, GU071348, GU071354, GU071355, GU071312, GU071314, GU071353, GU071316, GU071317, GU071318, GU071319, GU071284, GU071295, GU071294, GU071282, GU071298, GU071292, GU071631	Java, North Sumatra, South Sulawesi, West and East Sumba, West and East Flores, Alor, Ternate,
B8	GU071432, GU071561, GU071562, GU071567, GU071558, GU071552, GU071553, GU071445, GU071532, GU071619, GU071621, GU071622, GU071624, GU071625, GU071646, GU071648, GU071649, GU071650, GU071655, GU071656, GU071659, GU071661, GU071291	Java, West and North Sumatra, Mentawai, Lombok, West Flores, Alor

Table 4. continued

Genotype	Accession Number	Geographic origin / reported from
B9	GU071547, GU071574, GU071588, GU071591, GU071594, GU071589, GU071595, GU071527, GU071530, GU071528, GU071620, GU071672, GU071673, GU071675, GU071678, GU071679, GU071680, GU071681, GU071682, GU071683, GU071684, GU071685, GU071686, GU071687, GU071688, GU071689, GU071665, GU071667, GU071669, GU071708, GU071709, GU071710, GU071712, GU071713, GU071715, GU071716, GU071717, GU071691, GU071692, GU071693, GU071694, GU071697, GU071698, GU071699, GU071700, GU071701, GU071702, GU071705, GU071706, GU071707, GU071663, GU071364, GU071359, GU071642, GU071640, GU071645	West Sumatra, Nias, Mentawai, Lombok, West and East Sumba
C1	GU071419, GU071580, GU071581, GU071582, GU071583, GU071584, GU071538, GU071545, GU071551, GU071554, GU071556, GU071549, GU071557, GU071559, GU071564, GU071565, GU071568, GU071569, GU071570, GU071572, GU071576, GU071578, GU071578, GU071579, GU071541, GU071540, GU071542, GU071539, GU071537, GU071457, GU071435, GU071637, GU071361, GU071647, GU071652, GU071653, GU071654, GU071657, GU071658, GU071660, GU071315, GU071351, GU071329, GU071330, GU071333, GU071335, GU071289, GU071290, GU071297, GU071296, GU071629, GU071634, GU071299, GU071300, GU071301, GU071302, GU071303, GU071304, GU071306, GU071366, GU071397, GU071392, GU071393, GU071394, GU071395, GU071396, GU071385	Java, West Sumatra, South Sulawesi, East Sumba, West and East Flores, Alor, Ternate, Moluccas
C2	GU071550, GU071543, GU071544, GU071560, GU071566, GU071587, GU071585, GU071586, GU071446, GU071452, GU071449, GU071470, GU071475, GU071500, GU071501, GU071503, GU071506, GU071511, GU071461, GU071307, GU071525, GU071524, GU071639, GU071362, GU071360, GU071641, GU071323, GU071325, GU071327, GU071328, GU071332, GU071288, GU071293, GU071286, GU071608, GU071599, GU071611, GU071602, GU071597, GU071381, GU071386	Java, North and South Sumatra, South Sulawesi, East Sumba, West and East Flores, Alor, Papua
C5	GU071536, GU071534, GU071535, GU071533, GU071630, GU071632, GU071633,	North Sulawesi, Ternate
C6	GU071596, GU071606, GU071601, GU071603, GU071605, GU071607, GU071604	Papua
D1	GU071518, GU071635, GU071609	South Sulawesi, Papua
D3	GU071305, GU071600, GU071610, GU071598	Moluccas, Papua

The genotype and subgenotype of the 440 HBV isolates collected in the present study (Table 1) were determined from the sequence of the PreS2 region of the HBV genome, employing the genotype- and subgenotype-specific SNPs as described in Table 2.

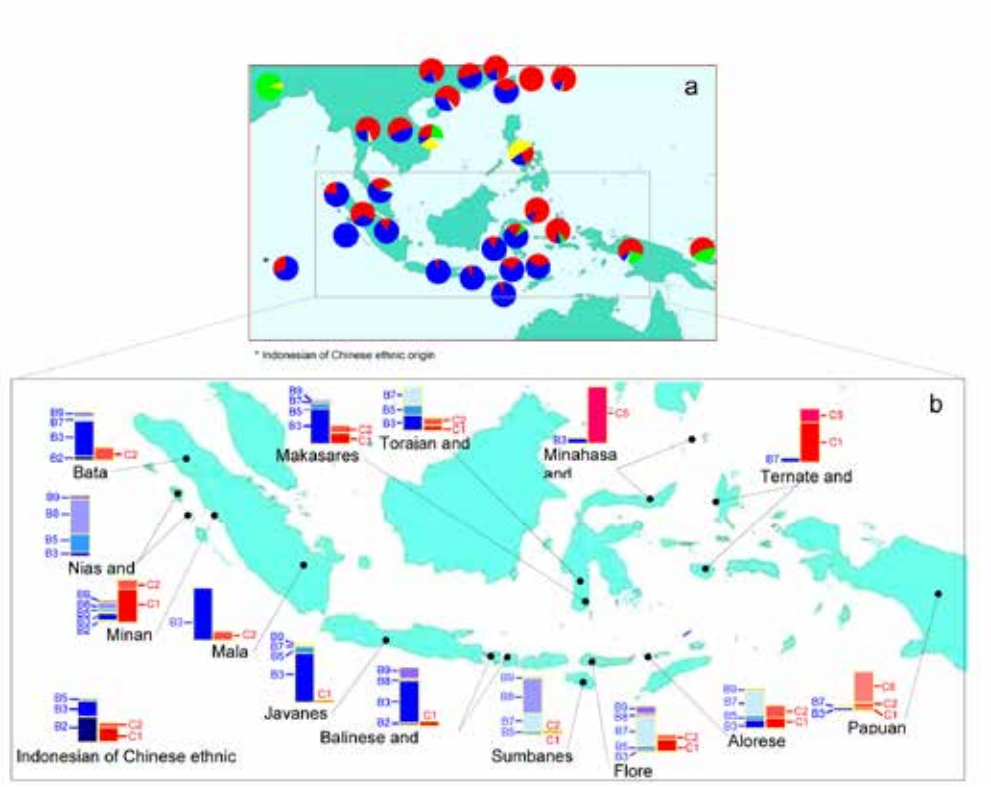


Figure 2. Ethnogeographical distribution of HBV genotypes/subgenotypes in the Indonesian archipelago. A total of 440 new HBV isolates were collected from 40 ethnic populations with a strict protocol to ensure the ethnic origins of their hosts to three previous generations (maternally and paternally) as described in Table 1. The isolates were genotypes/subgenotypes based on a set of diagnostic SNPs as shown in Table 2. Three genotypes, B, C and D, and their subgenotypes were determined as shown in Table 4. Data of previously published Indonesian HBV were added to the above, but only of isolates of well-defined ethnic origins (86) [1 (8), 34 (54), 55 (13), 57 (11)]. Figure 2a shows the distribution of HBV genotypes in Indonesia in comparison with isolates from mainland Asia and Oceania derived from published data (3691) [41 (100), 39 (720), 50 (332), 18 (146), 38 (367), 53 (382), 19 (209), 59 (776), 56 (211), 5 (220), 60 (67), 9 (62), 23 (51), 16 (48)]; note that these data were from isolates with defined geographical but not ethnic origins. Figure 2b shows details of HBV/B and HBV/C subgenotypes which are the main HBV genotypes in Indonesia. The genotypes/subgenotypes are: A (yellow), B (shades of blue), C (shades of red) and D (green).

Following the above scenario, HBV/C, shown to be dominant in populations of mainland Asia and in indigenous populations of Papua and Australia (Figure 2), but relatively low in the Austronesian-speaking populations, would have probably been introduced by the initial peopling of the archipelago. C1, which is the predominant HBV/C subgenotype in Indonesia, is also most prevalent in southern China [56]. The arrival of the Austronesian-speaking populations in the archipelago 3,000 to 5,000 yBP presumably displaced most HBV/C with the introduction of the Austronesian associated HBV/B. The observation of the different spectrum of HBV subgenotypes associated with WMP, CMP and SHWNG branches of the Austronesian-speaking population further support the suggestion of co-migration of HBV/B and its human hosts. And that the transmission of HBV in the distance past was mainly vertical from mother to her children, mimicking the transmission of the maternally inherited human mitochondrial DNA [34].

The other interesting finding in this study was the observation that HBV/B2, which is characteristic to the Chinese populations of mainland Asia and Taiwan [22, 32], was dominant also in Indonesians of Chinese ethnic origin, consistent with our previous observation [34]. Significantly, HBV/B3 was found to be the second major HBV/B subgenotype in the Indonesian Chinese. HBV/B3 has never been reported in populations of China and Taiwan. Thus, this observation presumably reflects the social interactions between indigenous and Chinese populations of Indonesia.

Several interesting deviations to the general pattern were observed, such as in the Austronesian WMP speaking populations of Minang of west Sumatra, the Mandar, Kajang and Toraja of south Sulawesi, and the mixed Austronesian-Papuan populations of Alor in east Nusa Tenggara. Some of these deviations could be traced to more recent population interactions and movements within the archipelago.

Independent of the speculations on its origin, the finding of the specific association of HBV subgenotypes with the ethnic populations of the Indonesian archipelago is of epidemiological and medical relevance. Study of mutations that underlie beta thalassemia in Indonesia, for example, has also indicated similar ethnic populations association in the distribution of some 30 beta-globin mutations [44]. The distribution of many other diseases in the Indonesian archipelago would probably be determined to various degrees by the genetic clustering of its ethnic populations.

Acknowledgements

The authors are grateful to Dr. Julius and colleagues, Department of Internal Medicine, Faculty of Medicine, Andalas University, Padang, Indonesia, Dr Soeyata and colleagues, Department of Internal Medicine, Faculty of Medicine, Sriwijaya University, Palembang, Indonesia, and Dr Zein Patiiha, Department of Medicine, Ternate General Hospital, Ternate, Indonesia, for their invaluable contribution in blood samples to this study.

References

1. Achwan WA, Muttaqin Z, Zakaria E, Depamede SA, Mulyanto, Sumohardjo S, Tsuda F, Takahashi K, Abe N, Mishiro S (2007). Epidemiology of hepatitis B, C, and E viruses and human immunodeficiency virus infections in Tahuna, Sangihe-Talaud Archipelago, Indonesia. Intervirol 50:408-411.
2. Arauz-Ruiz P, Norder H, Robertson BH, Magnius LO (2002). Genotype H: a new Amerindian genotype of hepatitis B virus revealed in Central America. J Gen Virol 83:2059-2073.
3. Bellwood P (1995). Austronesian prehistory in Southeast Asia: homeland, expansion and transformation. In: Bellwood P, Fox J J, Tryon D (eds). The Austronesians. The Australian National University, Canberra, pp 96-111.
4. Bellwood P (1997). Recent Indo-Malaysian prehistory: according to the languages. In: Prehistory of the Indo-Malaysian archipelago, 2nd edn. University of Hawai'i Press, Honolulu, pp 96-127.
5. Ding X, Mizokami M, Yao G, Xu B, Orito E, Ueda R, Nakanishi M (2001). Hepatitis B virus genotype distribution among chronic hepatitis B virus carriers in Shanghai, China. Intervirology 44:43-47.
6. Ethnologue. Irian Jaya (2000). In: Grimes B F, Grimes J E (eds). Ethnologue, languages of the world. Volume 1, 14th ed. SIL International, Dallas, pp 469-486.
7. Ethnologue. Papua New Guinea (2000). In: Grimes B F, Grimes J E (eds). Ethnologue, languages of the world. Volume 1, 14th edn. SIL International, Dallas, pp 741-799.
8. Fares MA, Holmes EC (2002). A revised evolutionary history of hepatitis B virus (HBV). J Mol Evol 54:807-814.
9. Gandhe SS, Chadha MS, Arankalle VA (2003). Hepatitis B virus genotypes and serotypes in Western India: lack of clinical significance. J Med Virol 69:324-330.
10. Hawkers RA, Boughton CR, Ferguson V, Vale TG (1981). The seroepidemiology of hepatitis in Papua New Guinea. II. A long-term study of hepatitis B. Am J Epidemiol 114 (4):563-573.
11. Higham CFW (2000). Prehistory, language and human biology: is there a consensus in East and Southeast Asia? In: Jin L, Seielstad M, Xiao C (eds).

Genetic, linguistic and archaeological perspectives on human diversity in Southeast Asia. World Scientific Publishing, Singapore, pp 3-15.

12. Huang AS, Tan CYH (2010). Achieving scientific eminence within Asia. Science 329:1471-1472.

13. Huy TT, Ushijima H, Quang VX, Luengrojanakul P, Kikuchi K, Sata T, Abe K (2004). Genotype C of hepatitis B virus can be classified into at least two subgroups. J.Gen Virol 85:283-292.

14. Huy TTT, Ngoc TT, Abe K (2008). New complex recombination genotype of hepatitis B virus identified in Vietnam. J Virol 82:5657-5663.

15. Ie SI, Thedja MD, Roni M, Muljono DH (2010). Prediction of conformational changes by single mutation in the hepatitis B virus surface antigen (HBsAg) identified in HBsAg-negative blood donors. Virol J 7:326.

16. Jazayeri MS, Basuni AA, Cooksley G, Locarnini S, Carman WF (2004). Hepatitis B virus genotypes, core gene variability and ethnicity in the Pacific region. J Hepatol 41:139-146.

17. Kao JH, Chen PJ, Lai MY, Chen DH (2000). Hepatitis B genotypes correlate with clinical outcome in patients with chronic hepatitis B. Gastroenterol 118:554-559.

18. Kao JH, Chen PJ, Lai MY, Chen DS (2004). Hepatitis B virus genotypes and spontaneous hepatitis B e antigen seroconversion in Taiwanese hepatitis B carriers. J Med Virol 72:363-369.

19. Kim H, Jee YM, Song BC, Shin JW, Yang SH, Mun HS, Kim HJ, Oh EJ, Yoon JH, Kim YJ, Lee HS, Hwang ES, Cha CY, Kook YH, Kim BJ (2007). Molecular epidemiology of hepatitis B virus (HBV) genotypes and serotypes in patients with chronic HBV infection in Korea. Intervirology 50:52-57.

20. Kim HJ, Lee DH, Gwak GY, Choi MS, Lee JH,Koh KC, Paik SW, Yoo BC (2007). Analysis of the core gene of hepatitis B virus in Korean patients. Liver International 27:633-638.

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

22. Kramvis A, Arakawa K, Yu MC, Nogueira R, Stram DO, Kew MC (2008). Relationship of serological subtype, basic core promoter and precore mutations to genotypes/subgenotypes of hepatitis B virus. J Med Virol 80:27-46.

23. Lim CK, Tan JTM, Khoo JBS, Ravichandran A, Low HM, Chan YC, Ton SH (2006). Correlations of HBV genotypes, mutations affecting HBeAg expression and HBeAg/Anti-HBe status in HBV carriers. Int J Med Sci 3:14-20.

24. Lusida MI, Nugrahaputra VE, Soetjipto R, Handajani M, Nagano-Fujii M, Sasayama T, Utsumi T, Hotta H (2008). Novel subgenotypes of hepatitis B virus genotypes C and D in Papua, Indonesia. J Clin Microbiol 46:2160-2166.

25. Magnius LO, Norder (1995). Subtypes, genotypes and molecular epidemiology of the hepatitis B virus reflected by sequence variability of the S-gene. Intervirology 38:24-34.

26. Marzuki S, Sudoyo H, Suryadi H, Setianingsih I, Pramoonjago P. Human genome diversity and disease on the island Southeast Asia (2003). In: Marzuki S, Verhoef J,

Snippe H (eds). Tropical diseases: from molecule to bedside, Adv Exp Med Biol 531, 3-18. Kluwer Academic/Plenum Publishers, New York, pp 3-18.

27. Miyakawa Y, Mizokami M (2003). Classifying hepatitis B virus genotypes. Intervirology 46:329-338.

28. Mulyanto, Depamede SN, Surayah K, Tsuda F, Ichiyama K, Takahashi, M, Okamoto H (2009). A nationwide molecular epidemiological study on hepatitis B virus in Indonesia: identification of two novel subgenotypes, B8 and C7. Arch Virol 154 (7):1047-1059.

29. Nagasaki F, Niitsuma H, Cervantes JG, Chiba M, Hong S, Ojima T, Ueno Y, Bondoe E, Kobayashi K, Ishii M, Shimosegawa M (2006). Analysis of the entire nucleotide sequence of hepatitis B virus genotype B in the Philippines reveals a new subgenotype of genotype B. J Gen Virol 87:1175-1180.

30. Naumann H, Schaefer S, Yoshida CF, Gaspar AM, Repp R, Gerlich WH (1993). Identification of a new hepatitis B virus (HBV) genotype from Brazil that expresses HBV surface antigen subtype *adw4*. J Gen Virol 74:1627-1632.

31. Norder H, Hammas B, Lufdahl S, Courouce AM, Magnius LO (1992). Comparison of the amino acid sequences of nine different serotypes of hepatitis B surface antigen and genomic classification of corresponding hepatitis B virus strains. J Gen Virol 73:1201-1208.

32. Norder H, Courouce AM, Coursaget P, Echevarria JM, Lee SD, Mushahwar IK, Robertson BH, Locarnini S, Magnius LO (2004). Genetic diversity of hepatitis B virus strains derived worldwide: genotypes, subgenotypes, and HBsAg subtypes. Intervirology 47:289-309.

33. Normile D (2009). SNP study supports southern migration route to Asia. Science 326:1470.

34. Nurainy N, Muljono DH, Sudoyo H, Marzuki S (2008). Genetic study of hepatitis B virus in Indonesia reveals a new subgenotype of genotype B in east Nusa Tenggara. Arch Virol 153:1057-1065.

35. Okamoto H, Imai M, Shimozaki M, Hoshi Y, Iizuka H, Gotanda T, Tsuda F, Miyakawa Y, Mayumi M (1986). Nucleotide sequence of a cloned hepatitis B virus genome, subtype ayr: comparison with genomes of the other three subtypes. J Gen Virol 67:2305-2314.

36. Okamoto H, Tsuda F, Sakugawa H, Sastrosoewignjo RI, Imai M, Miyakama Y, Mayumi M (1988). Typing hepatitis B virus by homology in nucleotide sequence: comparison of surface antigen subtypes. J Gen Virol 69:2575-2583.

37. Okamoto H, Nishizawa T (1992). Non-B non-G hepatitis virus gene, polynucleotide, polypeptide, virion, method for separating virion, and method for detecting virus. http://www.freepatentsonline.com/ EP1010759.html (with permission).

38. Olinger CM, Jutavijittum P, Hüschen JM, Yousukh A, Samountry B, Thammavong T, Toriyama K, Muller CP (2008). Possible new hepatitis B virus genotype, Southeast Asia. Emerg Infect Dis 14:1777-1780.

39. Orito E, Mizokami M, Sakugawa H, Michitaka K, Ishikawa K, Ichida T, Okanoue T, Yotsuyagani H, Lino S (2001). A case-control study for clinical and molecular biological differences between hepatitis B viruses of genotypes B and C.

Hepatology 33:218-223.

40. Ronquist F, Huelsenbeck JP (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572-1574.

41. Sakamoto T, Tanaka Y, Orito E, Co J, Clavio, Sugauchi F, Ozasa A, Quino A, Ueda R, Sollano J, Mizokami M (2006). Novel subtypes (subgenotypes) of hepatitis B virus genotypes B and C among chronic liver disease patients in the Philippines. J Gen Virol 87:1873-1882.

42. Sakamoto T, Tanaka Y, Simonetti J, Osiowy C, Berresen ML, Koch A, Kurbanov F, Sugiyama M, Minuk GY, McMahon BJ, Joh T, Mizokami M (2007). Classification of hepatitis B virus genotype into 2 major types based on characterization of a novel subgenotype in Arctic indigenous populations. J Infect Dis 196:1487-1492.

43. Sanders R, C, Lewis D, Dyke T, Alpers MP (1992). Marker of hepatitis B infection in Tari district, Southern Highlands Province, Papua New Guinea. P N G Med J 35:197-201.

44. Setianingsih I, Williamson R, Marzuki S, Harahap A, Tamam M, Forrest S (1998). Molecular basis of ß-thalassemia in Indonesia: application to prenatal diagnosis. Mol Diagn 3:11-20.

45. Stuvyer L, De Gendt S, Van Geyt C, Zoulim F, Fried M, Schinazi RF, Rossau R (2000). A new genotype of hepatitis B virus: complete genome and phylogenetic relatedness. J Gen Virol 81:67-74.

46. Sugauchi F, Mizokami M, Orito E, Ohno T, Kato H, Suzuki S, Kimura T, Ueda R, Butterworth LA, Cooksley WG (2001). A novel variant genotype C hepatitis B virus identified in isolates from Australian Aborigines: Complete genome sequence and phylogenetic relatedness. J Gen Virol 82:883-892.

47. Sugauchi F, Orito E, Ichida T, Kato H, Sakugawa H, Kakumu S, Ishida T, Chutaputti A, Lai CL, Ueda R, Miyakawa Y, Mizokami M (2002). Hepatitis B virus of genotype B with or without recombination with genotype C over the precore region plus the core gene. J Virol 76:5985-5992.

48. Sugiono D (2008). The mapping of languages in Indonesia. Language Center Press, Jakarta, pp 7.

49. Takahashi K, Akahane Y, Hino K, Ohta Y, Mishiro S (1998). Hepatitis B virus genomic sequence in the circulation of hepatocellular carcinoma patients: comparative analysis of 40 full-length isolates. Arch Virol 143:2313-2326.

50. Tangkijvanich P, Mahachai P, Komolmit P, Fongsarun J, Theamboonlers A, Poovorawan Y (2005). Hepatitis B virus genotypes and hepatocellular carcinoma in Thailand. World J Gastroenterol 11:2238-2243.

51. Tatematsu K, Tanaka Y, Kurbanov F, Sugauchi F, Mano S, Maeshiro T, Nakayoshi T, Wakuta M, Miyakawa Y, Mizokami M (2009). A genetic variant of hepatitis B virus divergent from known human and ape genotypes isolated from a Japanese patient and provisionally assigned to new genotype J. J Virol 83 (20):10538-10547.

52. The HUGO Pan-Asian SNP Consortium (2009). Mapping human genetic diversity in Asia. Science 326:1541-1545.

53. Toan NL, Song LH, Kremsner PG, Duy DN, Binh VQ, Koeberlein B, Kaiser S, Kandolf R, Toressi J, Bock CT (2006). Impact of the hepatitis B virus genotype and genotype mixtures on the course of liver disease in Vietnam. Hepatology 43:1375-1384.

54. Tryon D (1995). Proto-Austronesian and the major Austronesian subgroups. In: Bellwood P, Fox JJ, Tryon D (eds). The Austronesians: historical and comparative perspectives. The Australian National University, Canberra, pp 17-35.

55. Utsumi T, Lusida MI, Yano Y, Nugrahaputra VE, Amin M, Juniastuti, Soetjipto, Hayashi Y, Hotta H (2009). Complete genome sequence and phylogenetic relatedness of hepatitis B virus isolates in Papua, Indonesia. J Clin Microbiol 47:1842-1847.

56. Wang Z, Hou J, Wen S, Tanaka Y, Cheng J, Kurbanov F, Wang L, Jiang J, Naoumov NV, Mizokami M, Qi Y (2007). Distribution and characteristics of hepatitis B virus genotype C subgenotypes in China. J Viral Hepat 14:426-434.

57. Wibawa ID, Suryadarma IG, Mulyanto, Tsuda F, Matsumoto Y, Ninomiya M, Takahashi M, Okamoto H (2007). Identification of genotype 4 hepatitis E virus strains from a patient with acute hepatitis E and farm pigs in Bali, Indonesia. J Med Virol 79:1138-1146.

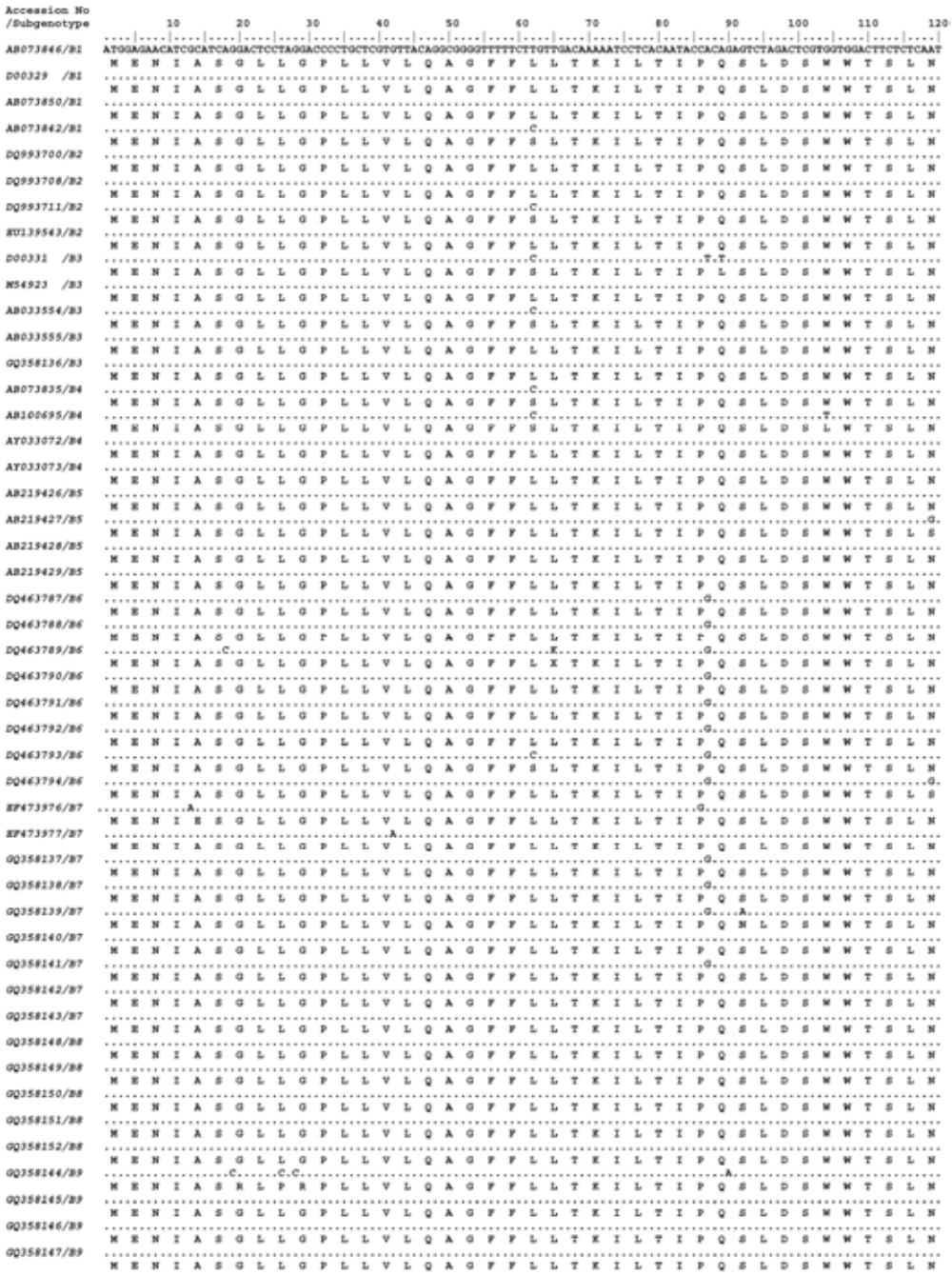
58. Yuasa R, Takahashi K, Dien BV, Binh NH, Morishita T, Sato K, Yamamoto N, Isomura S, Yoshioka K, Ishikawa T, Mishiro S, Kakumu S (2000). Properties of hepatitis B virus genome recovered from Vietnamese patients with fulminant hepatitis in comparison with those of acute hepatitis. J Med Virol 61:23-28.

59. Yuen MF, Sablon E, Tanaka Y, Kato T, Mizokami M, Doutreloigne J, Yuan HJ, Wong DKH, Sum SM, Lai CL (2004). Epidemiological study of hepatitis B virus genotypes, core promoter and precore mutations of chronic hepatitis B infection in Hong Kong. J Hepatol 41:119-125.

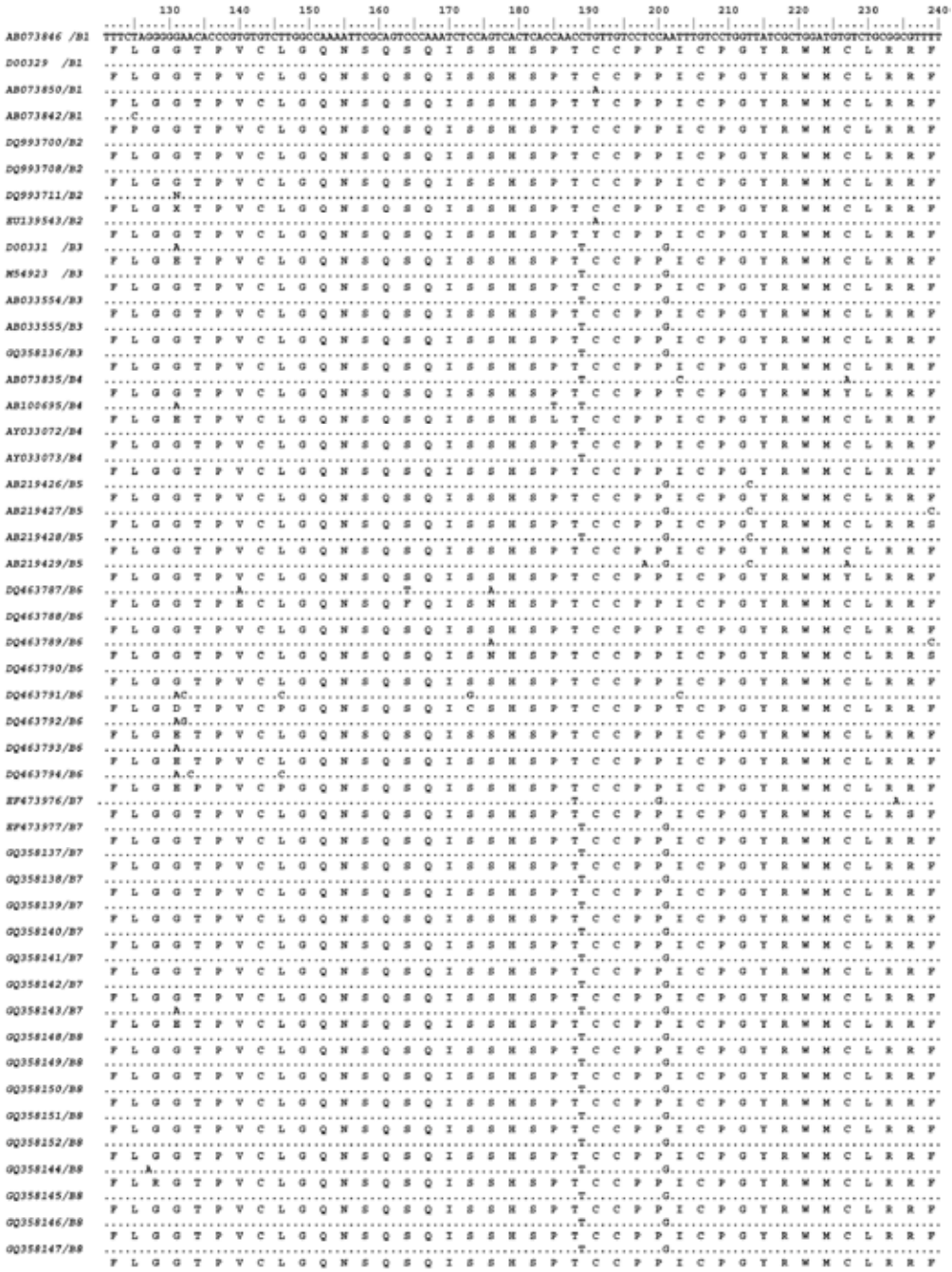
60. Zeng G, Wang Z, Wen S, Jiang J, Wang L, Cheng J, Tan D, Xiao F, Ma S, Li W, Luo K, Naoumov NV, Jou J (2005). Geographic distribution, virologic and clinical characteristics of hepatitis B virus genotype in China. J Viral Hepat 12:609-617.

61. Zhou Y, Holmes EC (2007). Bayesian estimates of the evolutionary rate and age of hepatitis B virus. J Mol Evol 65:197-205.

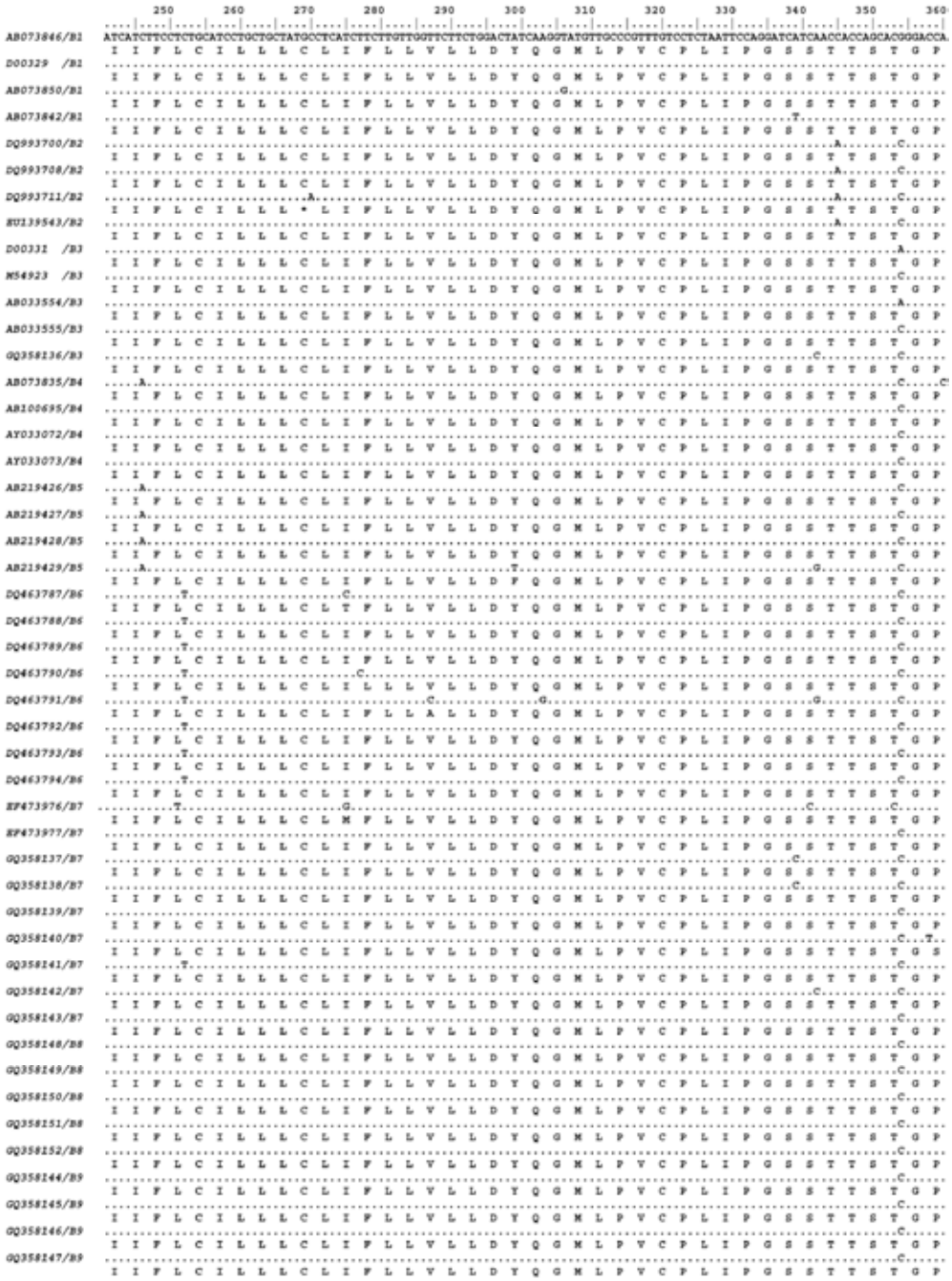
Supplementary Figure 1



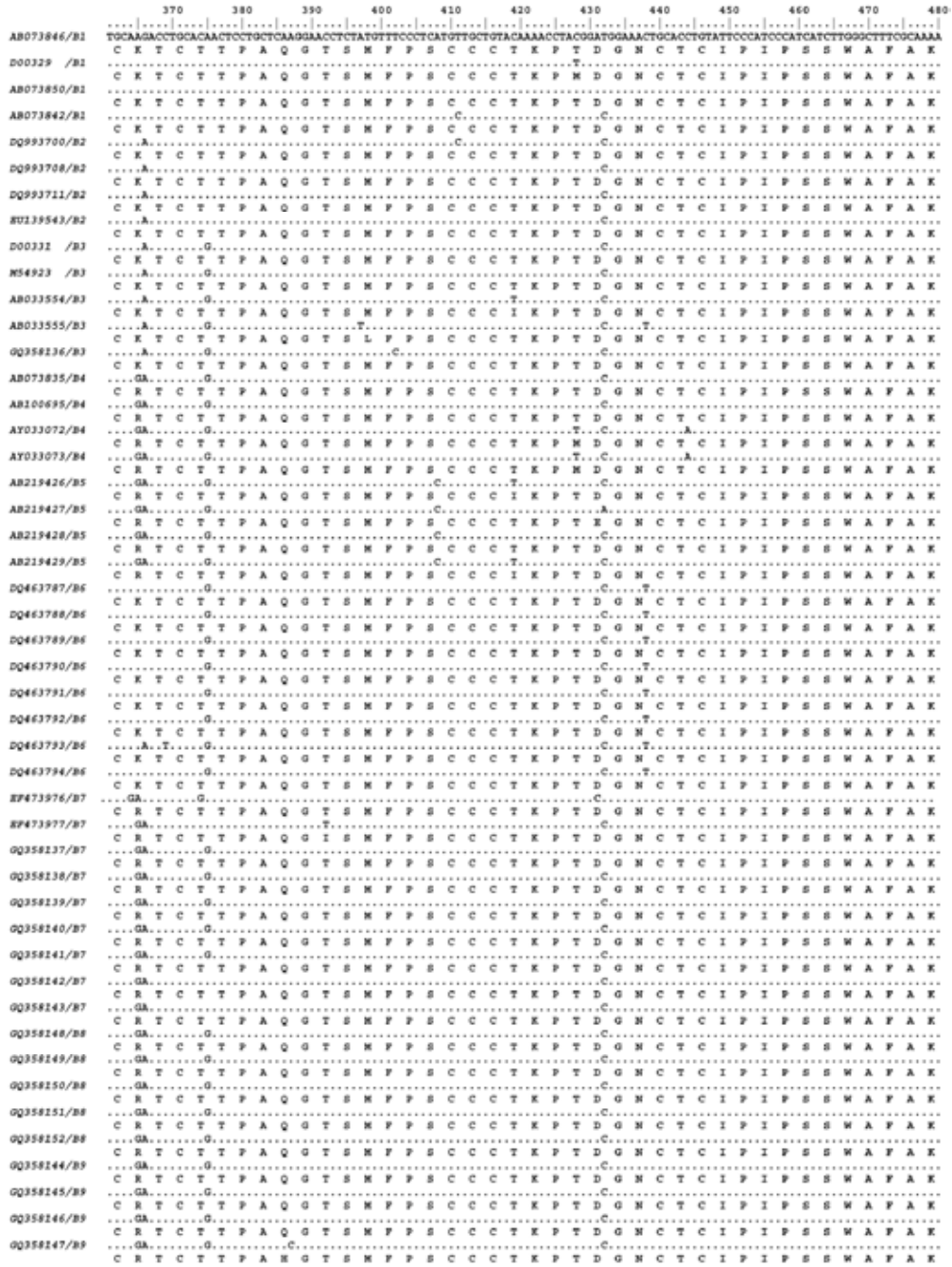
Supplementary Figure 1. continued



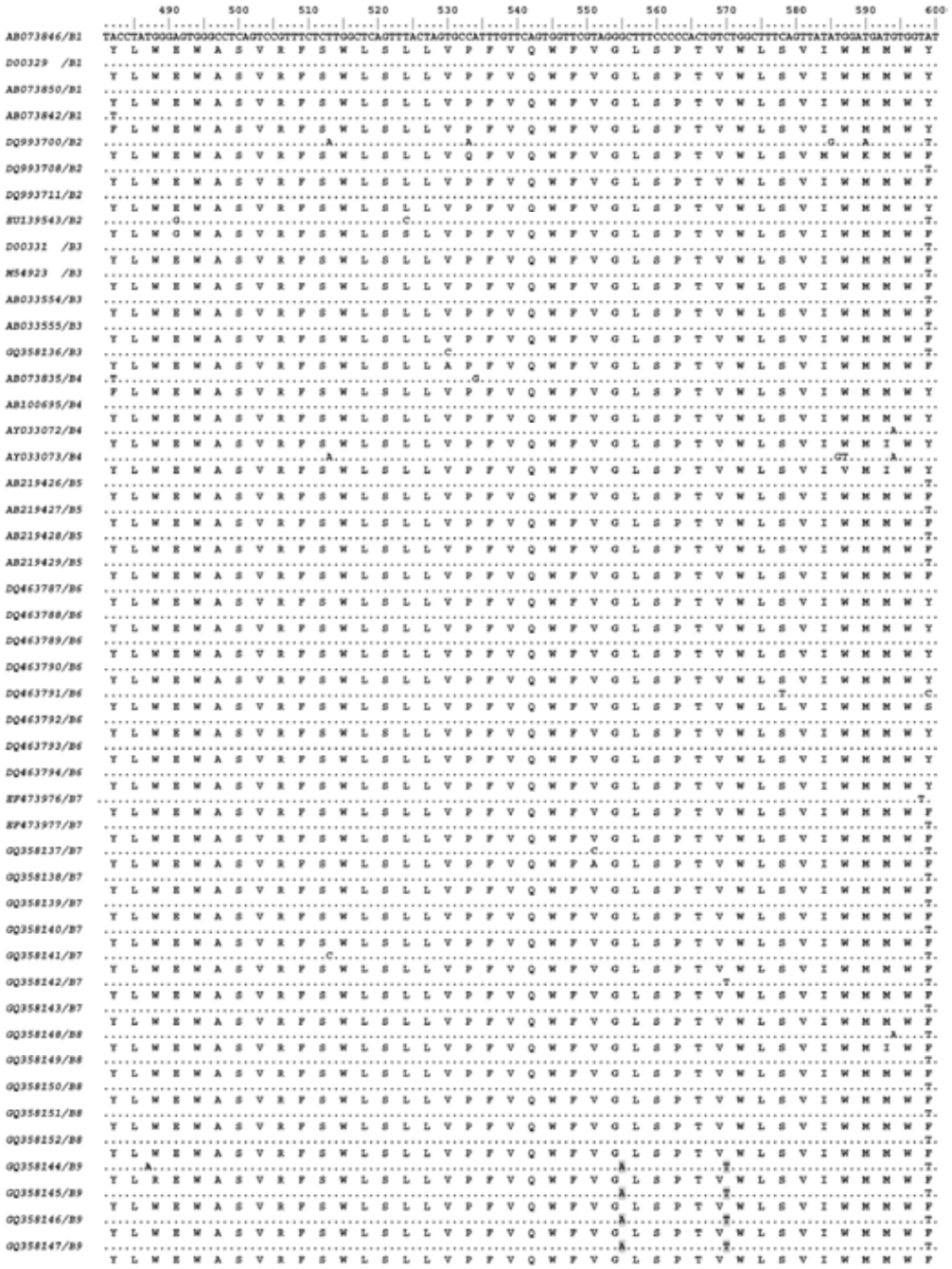
Supplementary Figure 1. continued



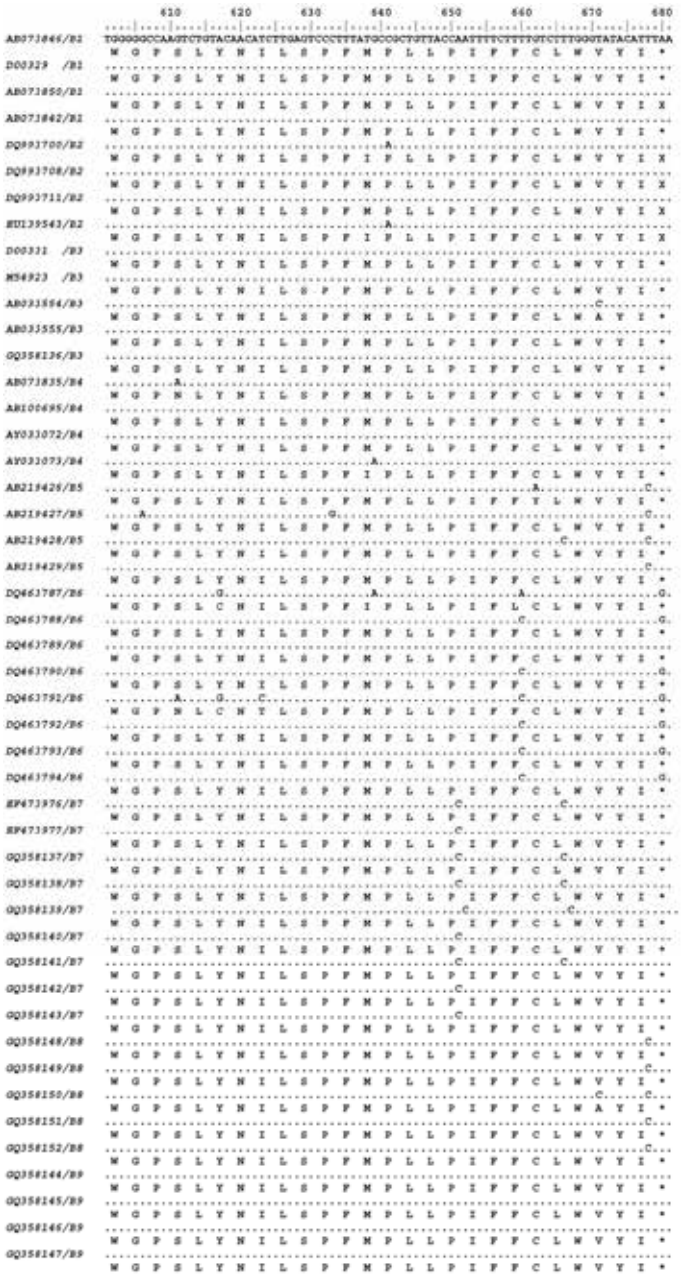
Supplementary Figure 1. continued



Supplementary Figure 1. continued

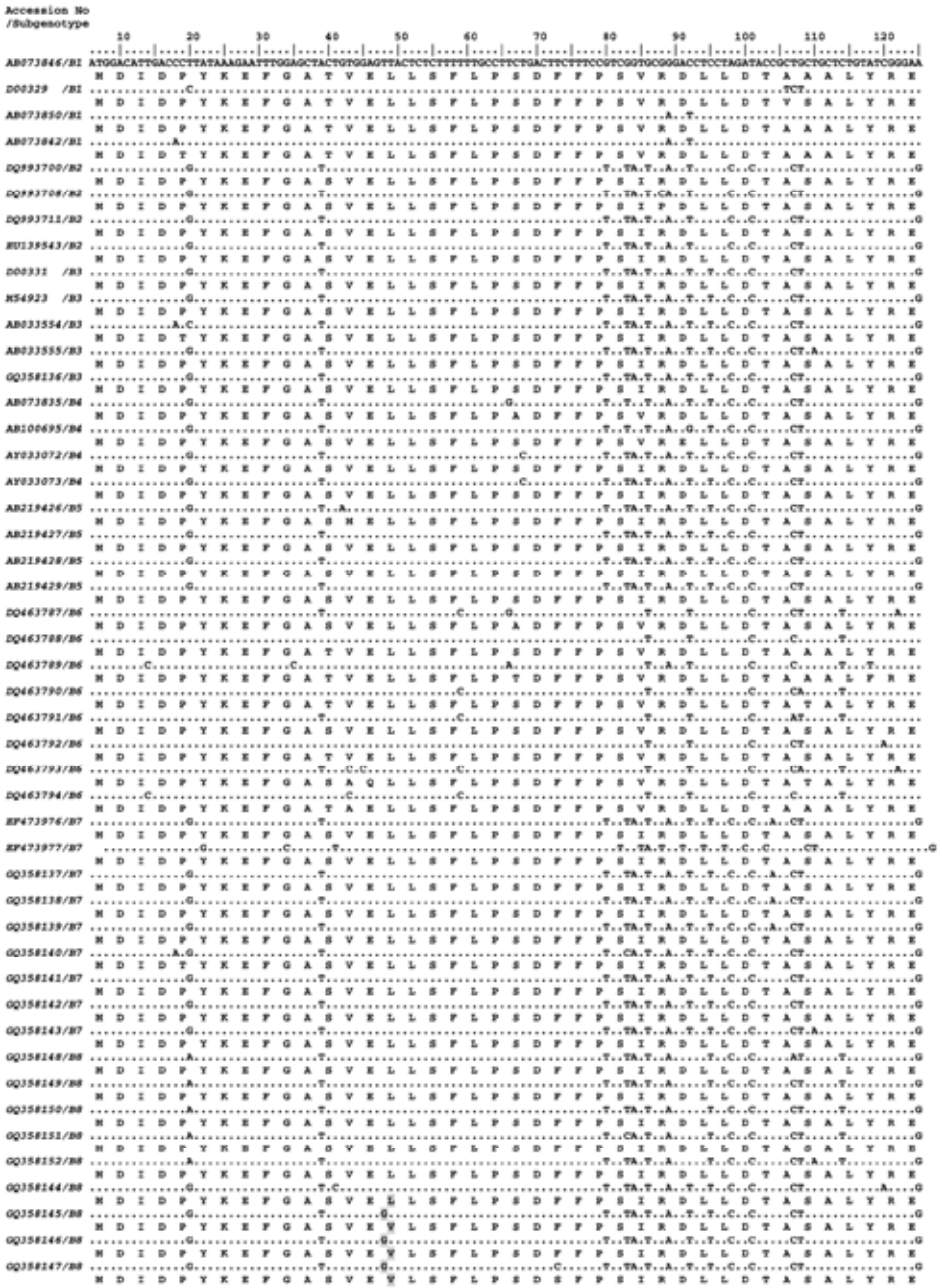


Supplementary Figure 1. continued

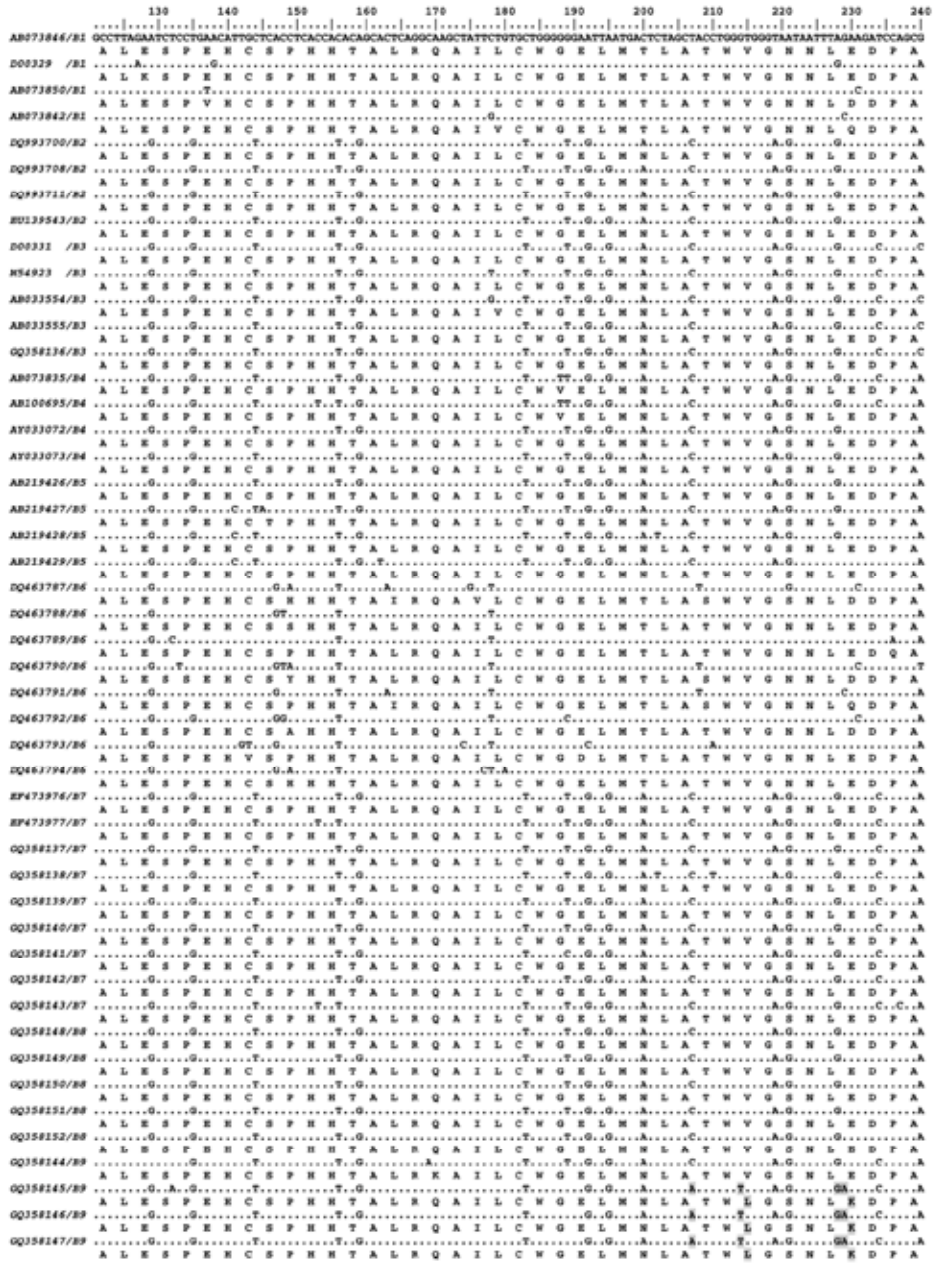


Nucleotide and amino acid configuration of HBV/B1–9 small surface sequences. Compared to other HBV/B subgenotypes, two nucleotide substitutions (G555A and C570T) not present in other/B subgenotypes were detected in HBV/B9. Nucleotide and amino acid numbering was based on the start codon of the surface sequence.

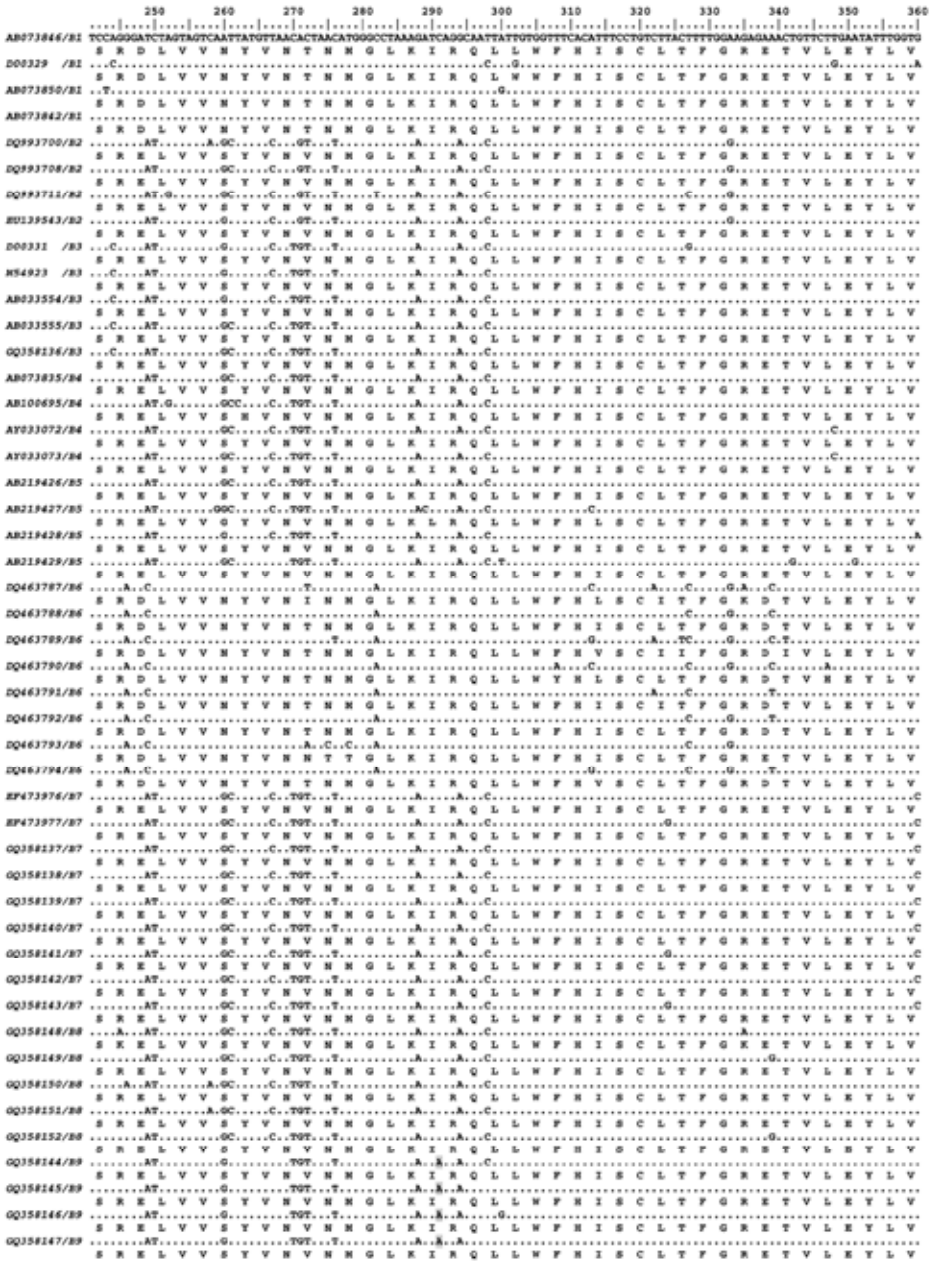
Supplementary Figure 2



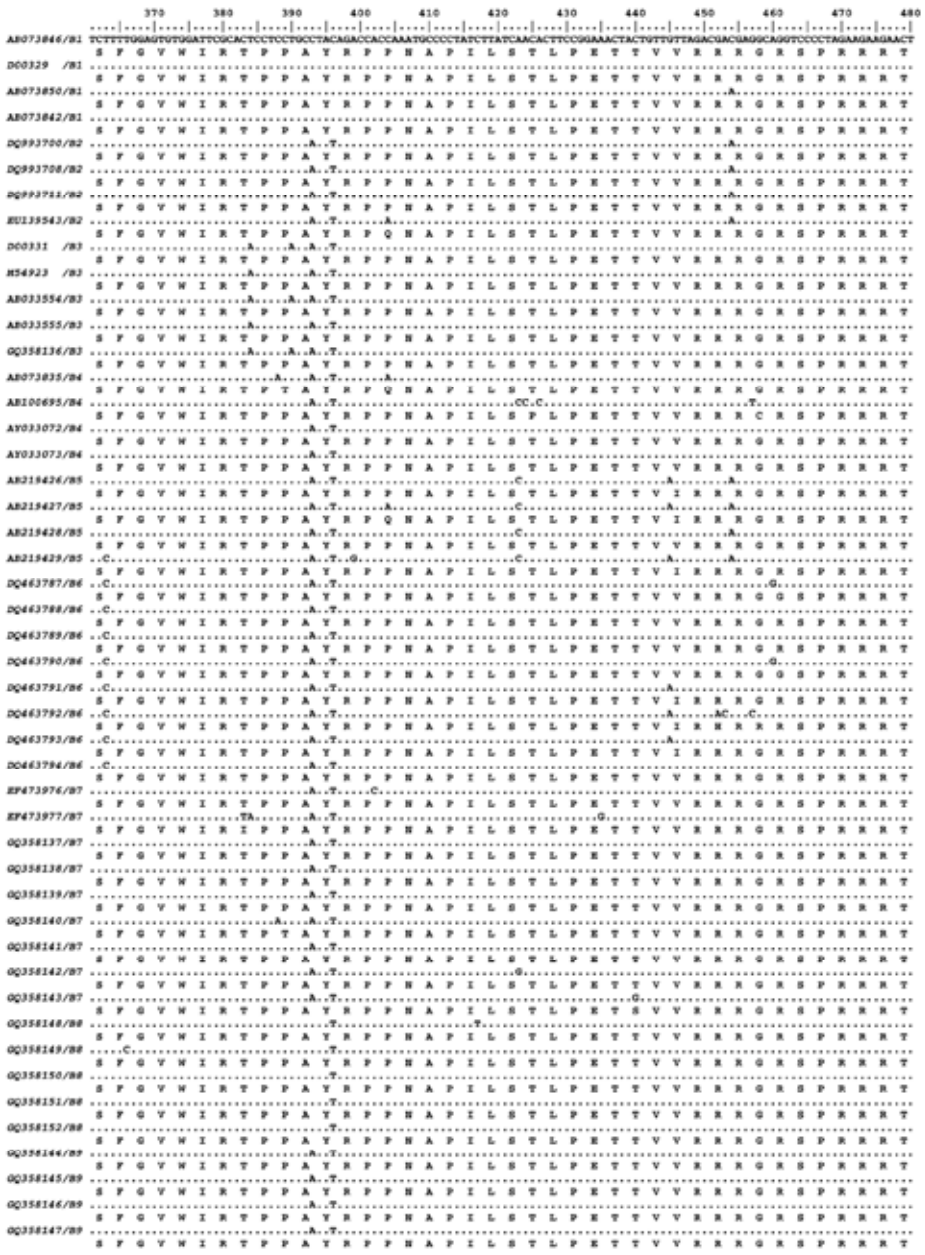
Supplementary Figure 2. continued



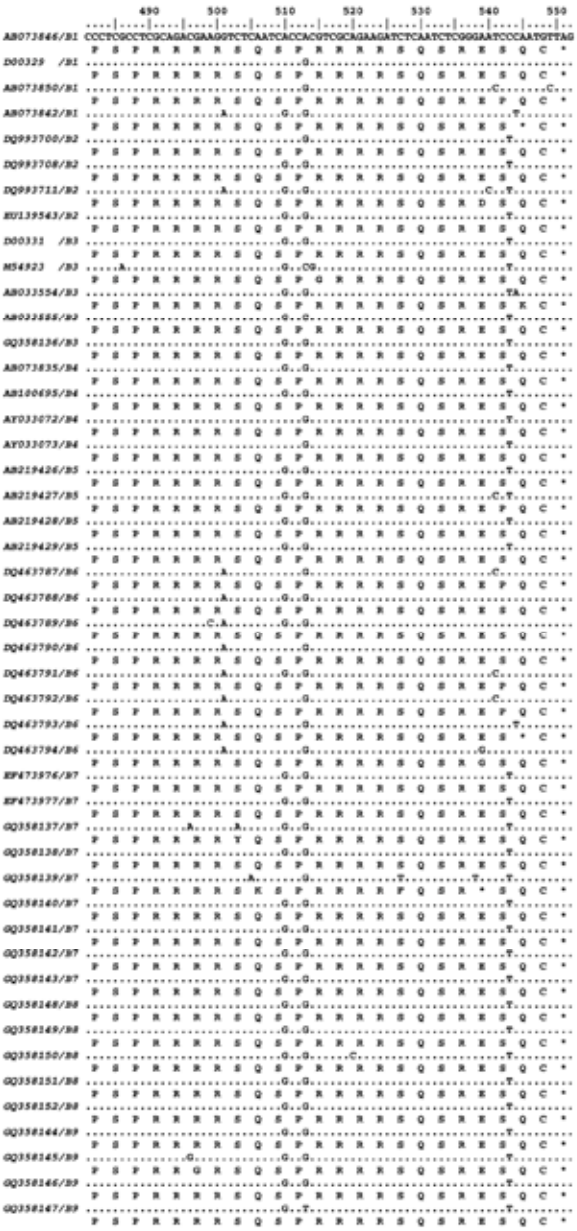
Supplementary Figure 2. continued



Supplementary Figure 2. continued



Supplementary Figure 2. continued



Nucleotide and amino acid configuration of HBV/B1–9 core sequences. Compared to other HBV/B subgenotypes, three nucleotide substitutions (T49G, G214T and G229A) causing amino acid substitutions (L15V, V72L and E77K, respectively) were detected in/B9, while 3 nucleotide substitutions (T207A, A228G, and C291A) caused silent mutation. These substitutions were not present in other HBV/B subgenotypes. Nucleotide and amino acid numbering was based on the start codon of the core sequences.

CHAPTER 3

Genogeography and immune epitope characteristics of hepatitis B virus genotype C reveals two distinct types: Asian and Papua-Pacific

Meta D. Thedja[1,2], David H. Muljono[1], Erick Sidarta[1], Turyadi[1], Susan I. Ie[1], Martono Roni[1],Jan Verhoef[2] and Sangkot Marzuki[1,3]

[1]Eijkman Institute for Molecular Biology, Jl. Diponegoro 69, Jakarta10430, Indonesia

[2]Eijkman Winkler Institute, Utrecht Medical Centre, Utrecht, The Netherlands

[3]Department of Medicine, Monash Medical Centre, Monash University, Clayton, Vic 3168, Australia.

Submitted

Summary

Distribution of hepatitis B virus (HBV) genotypes/subgenotypes is geographically and ethnologically specific. In the Southeast and East Asia and the Pacific regions, HBV genotype C (HBV/C) is prevalent with high genome variability, reflected by the presence of 13 of currently existing 16 subgenotypes in the Indonesian archipelago. We investigated the association between the genetic characteristics and ethnogeographical distribution of HBV/C various subgenotypes from Asia and Pacific, and further analyzed their immune epitopes within the core (HBcAg) and surface (HBsAg) proteins. Phylogenetic analysis of HBV/C complete sequences revealed the presence of two major groups, one for isolates from Southeast and East Asia (C1, C2, C5, C7, C8, C9, C10, and C14), and the other for those from Papua and Pacific (C3, C6, C11, C12, C13, C15, and C16). Analysis of HBcAg immune epitopes identified a single substitution (I59V) distinguishing the East and Southeast Asian isolates from those of the Papua-Pacific. Examination of HBsAg immune epitopes also showed two patterns of amino acid variation corresponding to the geographical origins of the isolates. Further characterization of HBsAg subtypic determinant revealed a west-to-east gradient with *adrq+* prominent in the Southeast-East Asia and *adrq-* in the Pacific. The *adrq-* indeterminate A159/A177 and a newly identified pattern of *adrq*-indeterminateV159/V177 were found in Papua and Papua New Guinea (PNG). This study indicates that HBV/C isolates can be classified into two types, the Asian and the Papua-Pacific, based on the virus genome diversity, immune epitopes, and geographical distribution, with Papua and PNG as the molecular evolutionary admixture region.

Introduction

Worldwide, an estimated two billion people have been infected with hepatitis B virus (HBV) and more than 240 million have chronic liver infections. About 600,000 people die every year due to the acute or chronic consequences of hepatitis B (WHO, 2012). In endemic region, such as Asia and Pacific where most individuals acquire the infection perinatally or in early childhood, up to 15-40% of individuals with chronic hepatitis B (CHB) will progress to cirrhosis, end-stage liver disease, or hepatocellular carcinoma (HCC) during their lifetime (Lavancy, 2004). HBV genetic variations, e.g. genotype and subtype, and mutations in some regions have been associated with different clinical manifestations such as development of cirrhosis and HCC, and response to treatment (Yang *et al.* 2008; Cao, 2009).

HBV has been classified into eight genotypes, A to H, defined by more than 8% sequence divergence over the entire genome (Okamoto *et al*., 1988). One additional HBV strain detected in Vietnam and Laos, and another identified in a patient in Japan were tentatively proposed as new genotypes I and J, respectively (Huy *et al*., 2008; Tatematsu *et al*., 2009). Based on some antigenic determinants of the surface antigen (HBsAg), nine serological types, referred to as subtypes—*adw2*, *adw4, adrq+, adrq-, ayw1, ayw2, ayw3, ayw4* and *ayr*—have been identified (Couroucé-Pauty *et al*., 1978; Norder *et al*., 1992). HBV genotypes and serotypes have a distinct geographical distribution worldwide, parallel and presumably evolve in populations of different ethnic origins (Norder *et al*., 2004).

In Asia and Pacific islands, HBV/B and HBV/C are the predominant genotypes. Compared to genotype B, genotype C is more often associated with higher rates of hepatitis B *e* antigen (HBeAg) carriers, lower rates of spontaneous HBeAg seroconversion, higher HBV DNA levels, with higher histological activities and higher proportion of patients developing cirrhosis and HCC (Kao *et al*., 2000; Orito *et al*., 2001; Zeng *et al*., 2005). In Indonesia, HBV/C is largely found in populations of the eastern islands, mostly in agreement with *adrq+* and *adrq*-indeterminate distribution (Lusida *et al*., 2008), while genotype B is typical in populations of the western islands of Indonesia, in parallel with the distribution of subtype *adw* (Sastrosoewignjo *et al*., 1991; Mulyanto *et al*., 1997).

Based on its genome diversity, HBV/C has been classified into sixteen subgenotypes, C1 to C16, each with specific geographical distribution. HBV/C1 (Cs) and C2 (Ce) were found predominantly in two different regions: C1 in Southeast Asia and C2 in east Asia (Chan *et al*., 2005; Norder *et al*., 2004; Thedja *et al*., 2011). HBV/C3 was found in the Oceania (Norder *et al*., 2004), C4 in Australian Aborigines (Sugauchi *et al*., 2001), with C5 and C7 in the Philippines (Cavinta *et al*., 2009; Sakamoto *et al*., 2006). Six other subgenotypes, C6, C8, C9, C10, C11, C12, and the recently reported C13, C14, C15, and C16 were found in the Indonesian archipelago (Mulyanto *et al*., 2009, 2011, 2012; Lusida *et al*., 2008). These ten subgenotypes were distinctly distributed: HBV/C6 in isolated populations of part of Papua, C8 in Nusa Tenggara and some western part of Indonesia (Denpasar, Jakarta, Banjarmasin, and Palembang), C9 in Timor Leste, and C10 in Nusa Tenggara, while C11-16 were found in Papua. This unique distribution pattern of HBV/C subgenotypes is of interest; thirteen (C1, C2, C5, C6, C8-16) of the sixteen existing HBV/C subgenotypes prevail in Indonesia, with some confined to certain parts of the archipelago. This situation is in contrast with mainland Asia, where only two subgenotypes (C1 and C2) are observed.

The genetic diversity of HBV has been suggested to be associated with natural selection influenced by host ethnic-related genetic background (Jazayeri *et al*., 2004a), reflected by divergence of amino acid substitutions within certain regions of HBV structural proteins, particularly the surface (HBsAg) and the core (HBcAg) antigens (Jazayeri & Carman *et al*., 2005). These two proteins are important because HBsAg contains T-cell and B-cell epitopes that define HBV variants (Carman *et al*.,1997; Ferrari *et al*., 1991; Tai *et al*., 1997), while HBcAg possesses immunologic targets of host immune response that determine the course of HBV infection (Jazayeri & Carman, 2005; Kim *et al*., 2007). Several Human Leukocyte Antigen (HLA)-restricted T cell epitopes within HBsAg and HBcAg have been proposed and different epitopes may present in consequence of the diverse distribution of HLA in populations in distinct geographical regions (Thursz *et al*., 2011).

Several studies on the association between genetic variation of HBV and the host have been reported (Bertoletti *et al*., 1994; Mohamadkhani *et al*., 2009; Thedja *et al*., 2011). The variation of HBV genetic characteristics has been extensively investigated for genotype B (Mulyanto *et al*., 2010; Thedja *et al*., 2011), but largely undefined for genotype C. Further, the knowledge on how the host-virus interaction shapes the molecular epidemiology pattern of HBV infection remains unclear. With ethnic diversity among the highest in the world, the Asia-Pacific region offers a unique host setting for HBV infection; its coincidence with the highly diverse distribution of HBV/C subgenotypes has never been studied. We carried out this study to investigate the association between HBV/C molecular characteristics and its ethnogeographical distribution, by examining various HBV/C subgenotype sequences from the Asia and Pacific region, with further analysis on the immune epitope characteristics of the core and surface proteins.

Results

Phylogenetic analysis of HBV complete genome sequences

Phylogenetic analysis based on 84 HBV complete sequences retrieved from GenBank confirmed the clustering of eight HBV genotypes and their subgenotypes as shown in Fig. 1. Interestingly, of 62 HBV/C isolates, two major clusters were observed: one of 35 isolates [C1 (10), C2 (9), C5 (7), C7 (1), C8 (4), C9 (1), C10 (1), and C14 (2)], and the other of 25 isolates [C3 (2), C6 (12), C11 (2), C12 (4), C13 (3), C15 (1), and C16 (1)]. The first cluster associated mainly with Southeast and East Asian countries, while the second cluster with those of Papua and Pacific region. The remaining two C4 isolates of

Northern Australia, however, belonged to a distinct lineage that was more distant compared to the two clusters.

Nucleotide divergence of HBV/C strains

Over the complete genome of 104 isolates (47 additional isolates retrieved from GenBank and 57 of 62 isolates used for phylogenetic tree construction), the nucleotide divergence between subgenotypes was higher than 4% (Table 1), with the exception of C2 to C14 (3.91%). The genetic divergence among HBV/C subgenotypes specific for Indonesia (C6, C8, C11, C12, C13, and C14) was as high as 5.92% for C12 to C13, while C6 to C11 had the lowest genetic divergence (4.02%). HBV/C1 and C2, which are specific for the Asian mainland, showed low genetic divergence (4.37%). Among all HBV/C isolates, C12 had the highest intra-subgenotype nucleotide divergence (3.25 ± 1.58%). The most deviating cluster, HBV/C4, had the highest evolution distinction to C1, C2, C3, C5, C6, C8, C11, C12, C13, and C14, supported by 7.16%, 6.36%, 6.01%, 7.71%, 6.57%, 6.26%, 6.19%, 6.69%, 7.44%, and 6.77% nucleotide divergence, respectively.

Evidence of recombination with HBV/B was detected by Bootscan analysis in the sixteen HBV/C subgenotypes, particularly in the precore to core region spanning from nt 1820 to nt 2350 with various length of recombination (Fig. S4).

Variation of amino acids within immunoepitopes of HBV/C strains

Inspection of HBsAg major class I HLA-A2-restricted Cytotoxic T Lymphocyte (CTL) epitopes (residues 20–28: FLLTRILTI) (Bertoletti *et al*., 1997) from 184 HBV/C sequences showed an R24K substitution in all HBV/C3 isolates of New Caledonia, C5 and C7 of the Philippines, as well as C8, C9, and C15 of Indonesia (Fig. 2). Substitutions were also observed at positions 44 and 47 located within a class I HLA-A2-restricted T-cell epitope (residues 41–49) (Tai *et al*., 1997). At position 44, a G44E substitution was identified in all HBV/C6, C11, C12, C13, C14, C15, C16, and unclassified HBV/C isolates from Papua New Guinea (PNG) and Tonga, with the majority of isolates from Fiji (18, 90%), Kiribati (3, 75%), and Vanuatu (8, 40%) also showed this pattern (Fig.S1). High amino acid variability was found at residue 47, particularly in Papua-Pacific HBV/C isolates (Fig.2 and Fig.S1). In the HBsAg B-cell epitopes of the *a* determinant (residues 124–148) (McMahon *et al*., 1992; Tiollais *et al*., 1985), an I126T substitution was only unanimously detected in HBV/C5 isolates from the Philippines, while P127T was identified in all HBV/C4 and 2 (10%) HBV/C isolates from Vanuatu (Fig.2 and Fig.S1). Within a class II HLA-DR-restricted T-cell epitope (residues 16-31) (Min *et al*., 1996;

Nayersina *et al*., 1993), a G18V switching was observed in all HBV/C6, C11, C12, C13, C15, C16, and majority of HBV/C isolates from PNG and from Pacific, but not in the other C subgenotypes. Furthermore, at residue 213, a substitution from L to I was detected in all isolates of HBV/C6, C7, C9, C11, C12, C13, C15, C16, and C from PNG and from Pacific, while HBV/C3, C4, and C8 showed similar prevalence for either L or I. Notably, most HBV/C6, C11, C12, C13, C15, and C16 isolates shared the same amino acid variation motifs with HBV/C isolates from Pacific with V18, E44, and I213, while isolates of the other HBV/C subgenotypes had G18, G44, and L213 motif. In comparison, the most distanced subgenotype, HBV/C4, showed more substitutions: G56, T68, S113, T114 (not shown), T127, A184, and I198.

Result of the analysis of 147 amino acid sequences of HBcAg for known T helper, CTL, and B cell recognition sites is shown in Fig.3 and Fig. S2. In the T helper epitopes (residues 1-20, 50-69, and 117-131), of the 81 HBV/C isolates from East and Southeast Asia, 6 isolates had variations within residues 1-20 with V13A as the most frequent substitution. In contrast, only one isolate from Fiji of Papua-Pacific had S12P and E14A substitutions (Fig.S2). Interestingly, within residues 50–69, a single substitution—I59V—distinguished the East and Southeast Asia isolates (HBV/C1, C2, C5, C7, C8, C9, and C10) from those of the Papua-Pacific (HBV/C3, C6, C11, C12, C13, C15, C16, and C Pacific). Notably, HBV/C14 that was phylogenetically grouped into the East and Southeast Asian cluster (Fig.1) had V59 that marked the Papua-Pacific strains (Fig.3 and Fig.S2). In residues 117–131, 13 of 75 isolates from East and Southeast Asia had P130T/L/Q/A substitution. The two most distinct HBV/C4 isolates from the Australian Aboriginals showed I59 and L/I130 with no variation within residues 1-20.

For HLA class-I-restricted epitopes in HBcAg (residues 18–27, 84–101, and 141–151) (Bertoletti *et al*., 1994; Ehata *et al*., 1991, 1992), most HBV/C isolates from Asia and Papua-Pacific had I27 within residues 18–27 (FLPSDFFPSI), except for 6 isolates from Asia (Japan, China, Hongkong, Myanmar, Vietnam, and Indonesia) that had V27. Various amino acid substitutions were observed in residues 84–101, of which V91 was consistently identified in isolates from East and Southeast Asia (C1, C2, C5, C7, C8, C9, and C10), Papua (C6, C11, C12, C13, C15, and C16), and Vanuatu of Pacific. However, HBV/C isolates from Polynesia (C3) and those from Fiji, Tonga, and Pacific had I91. The most divergent subgenotype, HBV/C4, had V91. Examination of HBcAg B-cell epitopes (residues 74–89, 107–118, 130–138, and 148–160) (Jazayeri & Carman, 2005) showed no specific variation in all HBV/C isolates.

HBsAg subtypes of HBV/C strains and their distribution

Amino acid variations that determine HBsAg subtypes of 271 HBV/C isolates (164 of the East and Southeast Asia, 107 of the Papua-Pacific, and 2 of North Australia) were identified (Table 2), including the combination of amino acids at positions 159 (A/V) and 177 (V/A) that specifies the subtype *adr* into *q+* and *q-* patterns (Norder *et al*., 1994). A clear signature of HBsAg *q* subdeterminant variability that separated the HBV/C isolates from Asia and from Papua-Pacific was observed.

The distribution of the subtypes with respect to their countries/geographical origins is illustrated in Fig. 4. Of 81 HBV/C1 isolates (Thailand 9, Vietnam 5, Myanmar 8, Malaysia 1, Indonesia 56, and China 2), the majority (60; 74.1%) could be classified into *adrq+* A159/V177 subtype, while 13 (16%), 6 (7.4%), and 1 (1.2%) belonged to *ayr*, *adw2* and *ayw1* subtypes, respectively. Interestingly, 1 (1.3%) strain from Thailand had Valine at both 159 and 177 residues. Since the unique V159/V177 combination has not been reported, we provisionally designate this pattern as another form of *adrq*-indeterminate, in addition to the previously reported *adrq*-indeterminate A159/A177 combination (Lusida *et al.*, 2008). Likewise, of 61 HBV/C2 isolates (China 14, Japan 21, Korea 3, and Indonesia 23), most (54; 88.5%) could be classified into *adrq+*, while the rest into *adw2* (4; 6.6%), *ayw1* (1; 1.6%), and *adrq*-indeterminate A159/A177 (2; 3.2%). All 13 HBV/C5 isolates from the Philippines and Indonesia belonged to *adw2* subtype, while isolates of HBV/C7 from the Philippines, C9 from Timor Leste, and C8, C10 and C14 from Indonesia had *adrq+* subtype.

Of 18 HBV/C6 isolates from Papua, 9 belonged to *adrq+* and 9 to *adrq*-indeterminate A159/A177 subtypes, whereas 2 C11, 4 C12, 3 C13, 1 C15, and 1 C16 isolates were of *ayr*, *adrq+*, *adrq*-indeterminate A159/A177, *adrq+*, and *adrq*-indeterminate A159/A177 subtypes, respectively. Of 10 HBV/C isolates from PNG, 8 had *adrq+* and 2 a*drq*-indeterminate V159/V177 subtypes. Further, among 64 HBV/C isolates from Pacific region (Vanuatu, Fiji, Kiribati, and Tonga), the majority of isolates (51; 79.7%) had *adrq-* subtype, while 10 belonged to *adrq*-indeterminate V159/V177, 2 to *adrq+*, and 1 to *ayw1* subtypes. Two HBV/C3 isolates derived from the Pacific region had *adrq-* subtype. Two HBV/C4 isolates from Australian Aboriginals were distantly related to both the Asian and the Pacific groups, showing a pattern typical for *ayw3* subtype.

All the 87 S gene sequences generated in this study had been deposited in GenBank database with accession numbers JQ740646-JQ740732.

Discussion

Separation of HBV/C into two major types: Asian and Papua-Pacific

HBV genotype C has been known to be predominant in Asia and Pacific region in addition to genotype B (Miyakawa *et al*., 2003; Jazayeri *et al*., 2004; Norder *et al*., 2004; Kurbanov *et al*., 2010; Kao, 2011). Sixteen HBV/C subgenotypes have been identified, mostly in Asia, some in the Pacific region, and one in northern Australia. Result of the present study reveals that 15 of these subgenotypes clustered into two major types of HBV/C, the Asian and the Papua-Pacific, whereas one subgenotype (C4) was confined to Northern Australia and had distinct genome characteristics (Table 1). Phylogenetic analysis of HBV/C complete genome clearly demonstrated the separation of HBV/C subgenotypes of Asian (East and Southeast Asian) countries (C1, C2, C5, C7, C8, C9, C10, and C14) from those of Papua and Pacific region (C3, C6, C11, C12, C13, C15 and C16) as seen in Fig. 1. Another interesting finding regarding the Asian subgenotypes was noted. Consistent with previous reports, two subgenotypes (C1 and C2) were prominent in Southeast Asia and East Asia, respectively, with homogenous distribution (Chan *et al*., 2005; Huy *et al*., 2004; Norder *et al*., 2004). However, the other six subgenotypes (C5, C7, C8, C9, C10, and C14) were observed only in Southeast Asia with heterogeneous distribution.

Result of the examination of variants within the CTL immune recognition sites of HBV/C core region is consistent with the genetic separation, showing a clear division of Asian and Papua-Pacific patterns. Significantly, we discovered a critical polymorphism 59I/V within the 183 amino acids of HBcAg distinguishing HBV/C into the Asian and Papua-Pacific types, with the exception of C14 that has Papua-Pacific 59V characteristics (Fig. 3). This 59I/V polymorphism warrants further study since it is located in the highly immunogenic T helper epitopes (Ferrari *et al.,* 1991; Kim *et al*., 2007).

The amino acid sequence of the core region as the important target for immune-mediated viral clearance by T helper cells, B cells, and CTL response is relatively conserved (Chisari, 2000; Norder *et al*., 2004). However, variations within the core region have been observed and associated with its function in induction of host immune response (Jazayeri & Carman, 2005; Mohamadkhani *et al*., 2009; Pairan & Bruss, 2009). In this study, by using HBcAg specific motifs recognized in the Asian and Papua-Pacific populations (Jazayeri *et al*., 2004b, Carman *et al*., 2005), we discovered that the isolates from East and Southeast Asia had more amino acid substitutions compared to those from Papua-Pacific (Fig. 3). In keeping with the reported low divergence of HBV strains in Papua-Pacific (Jazayeri *et al*., 2004b), the higher conservation of HBcAg could be

attributable to the more homogeneous immunity exerted by hosts of genetically less diverse populations in this geographical region (Friedlander *et al.,* 2008). This specific HBV core amino acid variation could be a consequence of host-virus interaction that shape the HBV-specific T-cell repertoire, and partly influenced by parallel evolution of geographically separated HBV lineages in the Asia and Papua-Pacific (Jazayeri *et al*., 2004b; Tan *et al*., 2008). Population-based studies in both regions are needed to understand the role of host factors in the selection of HBV strains as well as its clinical and public health implications.

The characteristics of immune epitopes of surface protein further support the segregation of HBV/C subgenotypes into the Asian and Papua-Pacific groups. We discovered three unique substitutions—G18V, G44E, and L213I— that separate the two groups (Fig. 2). Interestingly, two of these substitutions are located within important immune epitopes: G18V in the class II HLA-DR-restricted T-cell epitope (residues 16-31) and G44E in the class I HLA-A2-restricted T-cell epitope (residues 41-49). The HLA-DR-restricted epitope (residues 16-31) has been shown to have lower capacity in eliciting antibody response to HBsAg vaccination (Min *et al.,*1996), while the HLA-A2-restricted epitope (residues 41-49) is located in the hotspot mutational domain of HBsAg (residues 25-51). It was hypothesized that this domain could contribute to the protective cellular immunity in natural infection with HBV, and mutations within this domain could predispose the hosts to chronic infection (Tai *et al.,* 1997). Future functional studies on these mutations would lead to a better understanding of the complex virus-host interactions.

Some anomalies were also observed in this study. Two HBV/C14 isolates (AB644283 and AB644284) recently found in Papua clustered together with strains from Asia (Fig. 1). These isolates have 59V of the core protein characteristic of the Papua-Pacific type (Fig. 3), but they have 18G of the surface protein that marks the Asian type (Fig 2). More isolates of this subgenotype are expected to explain this phenomenon, since the two isolates were derived from hosts with different ethnic backgrounds, i.e. Austronesian and non-Austronesian (Mulyanto *et al.,* 2012). The two C4 isolates from Aboriginal population of Australia showed a distinct cluster unclassifiable into either the Asian or Papua-Pacific types, suggesting that this subgenotype might have a molecular evolution different from the other C subgenotypes.

All the analyses above indicate the presence of two types of HBV/C subgenotypes: the Asian and Papua-Pacific, based on the genetic characteristics of complete genome, the immune recognition sites within core and surface proteins, and the pattern of HBsAg subtypes. This finding is consistent with the host ethnical and geographical association as proved by the linguistic evidence: HBV/C of Asian type with the Austronesian

speaking populations in the East and Southeast Asia, and HBV/C of Papua-Pacific type with the Papuan speaking populations in the Papua-Pacific region (Bellwood, 1997; Tryon, 1995). Taken together, this finding could suggest that the distribution of HBV/C isolates follows linearly the prehistorical human dispersal, in agreement with our previous report on HBV/B subgenotype distribution in Indonesia (Thedja *et al*., 2011).

Papua and Papua New Guinea as the HBV/C admixture region

A closer look at the HBsAg subtype distribution of HBV/C reveals a west-to-east gradient with *adrq+* prominent in Asia and *adrq-* in the Pacific. Among 164 HBV/C isolates from the East and Southeast Asian countries (China, Japan, Korea, Myanmar, Vietnam, Thailand, Philippines, Malaysia, Indonesia, and Timor Leste), the majority were *adrq+* (75%), followed by *adw2* (14.0%), *ayr* (7.9%), and *ayw1* (1.2%). The remaining subtypes were *adrq*-indeterminate A159/A177 (1.2%) and a novel *adrq*-indeterminate V159/V177 (0.6%). In the Indonesian archipelago, specific pattern of HBsAg subtype distribution was also observed, with predominance of *adrq+* spanning from the western part to Nusa Tenggara islands in the east. In contrast to East and Southeast Asia, among 64 HBV/C isolates from the Pacific region (Vanuatu, Fiji, Tonga, and Kiribati islands), *adrq-* was the most prevalent (79.7%), followed by *adrq*-indeterminate V159/V177 (15.6%), *adrq+* (3.1%), and *ayw1* (1.6%).

In Papua and PNG, the regions between Southeast Asia and the Pacific, both *adrq*-indeterminate forms prevail in addition to *adrq+* and *adrq-*: A159/A177 in Papua and V159/V177 in PNG. Subtype *adrq*-indeterminate A159/A177 represents 41.9% of HBV/C in Papua, while subtype *adrq*-indeterminate V159/V177 accounts for 15.6% of HBV/C in PNG**.** This specific distribution could show that Papua and PNG are the regions where the switching from *adrq+* (51.6% in Papua) to *adrq-* (79.7% in PNG) occurred, characterized by the presence of the two *adrq*-indeterminate as intermediate forms. The importance of these geographical regions as the transitional zone of past migratory events from Asia into Pacific (Deka *et al.,* 2001), followed by long standing isolation of these populations could be the background of this HBV/C genetic segregation in Papua and PNG.

Diversity of HBV/C in Indonesia

Positioned in the middle of Asia-Pacific, Indonesia with around 500 ethnic populations dispersed in thousands of island, has a unique distribution of HBV genotypes and subgenotypes (Nurainy *et al*., 2008; Mulyanto *et al*., 2010). We recently demonstrated the specific association of the distribution of HBV genotypes/subgenotypes with different

host ethnic populations and geographical background (Thedja *et al*., 2011). The presence of thirteen HBV/C subgenotypes (C1, C2, C5, C6, C8-C16) in Indonesia is remarkable as to date no other region has been reported to have such a high variety of HBV/C isolates. This fact demonstrates that Indonesia has a much greater HBV/C genome diversity than other regions in Asia, in contrast to East Asian countries that have a more homogenous distribution of HBV/C subgenotypes. This varying distribution was also seen across the Indonesian archipelago, from two (C1 and C2) in the west to more heterogeneous (C5, C6, C8-C16) subgenotypes in the east. The finding of putative recombination between HBV/C and part of HBV/B sequence in the precore/core region gave additional contribution to the diversity of the HBV/C genome characteristics. This specific distribution accentuates the uniqueness of HBV/C diversity in Indonesia, the same as that of HBV/B reported previously (Thedja *et al.,* 2011).

In conclusion, the present study indicates that HBV/C isolates can be classified into two types, the Asian and the Papua-Pacific, based on the virus genome diversity, immune epitopes, and geographical distribution, with molecular evolutionary admixture occurred in Papua and PNG. More HBV/C isolates could be expected from these regions, since the chance of having HBV/C genetic admixture is greater there. Further investigation in both scientific and public health perspectives on the relevance of the two types of HBV/C subgenotypes need to be undertaken, together with other genotypes and subgenotypes prevailing in this region. This information would provide insights into the development of management strategy and the design of diagnostic tools and vaccine for HBV infection for such a genetically diverse host population.

Materials and Methods

HBV complete genome sequences and genetic relatedness analysis

Eighty-four HBV complete genome sequences were retrieved from GenBank, including 62 HBV/C isolates: 37 [C1 (3), C2 (1), C5 (3), C6 (12), C8 (4), C10 (1), C11 (2), C12 (4), C13 (3), C14 (2), C15 (1), and C16 (1)] from various geographical regions and ethnic populations of the Indonesian archipelago (Lusida *et al*., 2008; Mulyanto *et al*., 2009, 2010, 2011, 2012; Thedja *et al*., 2011) and 25 [C1 (7), C2 (8), C3 (2), C4 (2), C5 (4), C7 (1), and C9 (1)] from other countries in the Asia (Korea, China, Japan, Myanmar, Thailand, Vietnam, Malaysia, Philippines, and Timor Leste), the Pacific (Polynesia and New Caledonia), and Northern Australia, together with 22 isolates representing HBV/A (2), HBV/B (9), HBV/D (6), E (1), HBV/F (2), G (1), and H (1).

The 84 HBV sequences were aligned using ClustalW software (http://www.ebi.ac.uk/ClustalW/) and confirmed by visual inspection. Phylogenetic tree was constructed by neighbor-joining method and genetic distance was calculated using Kimura two-parameter method available in the Mega4 program. HBV strain of woolly monkey hepatitis virus (AY226578) was used as outgroup. Bootstrapping with 1000 replicates was performed to ensure the reliability of the tree.

To define the magnitude of inter-genotype and intra-genotype differences, pairwise analysis of nucleotide divergence was performed for existing HBV/C subgenotypes. To increase data validity, 47 additional HBV/C complete sequences from the Asia and Pacific were searched and analyzed along with the initial 62 HBV/C isolates. Five subgenotypes (HBV/C7, C9, C10, C15, and C16) were not included since only single sequence was available for each subgenotype. Totally, 104 HBV/C isolates were used in this analysis. Further, the putative recombination between each of HBV/C1-16 subgenotypes with other seven HBV genotypes was assessed by Bootscan analysis in the SimPlot program (http://sray.med.som.jhmi.edu/RaySoft/SimPlot/). Analysis was done for the complete sequence, and 200 bp windows size, 20 bp step, and 1000 bootstrap were used in analysis.

Additional HBV/C sequences and sample preparation

To obtain a better understanding of HBV/C subgenotype distribution in the Indonesian archipelago, S gene sequences were generated from 87 HBV/C isolates of ethnically defined origins determined in our previous study: HBV/C1 (53), C2 (22), C5 (6) and C6 (6) (Thedja *et al*., 2011). The ethnic background of the individuals had been ascertained for at least three previous generations both maternally and paternally (Marzuki *et al*., 2003). These samples originated from the islands of Sumatra (36), Kalimantan (1), Sulawesi (9), Flores (19), Sumba (1), Alor (3), Ternate-North Moluccas (5), Ambon-South Moluccas (7), and Papua (6) of the Indonesian archipelago (named Papua hereafter) (Fig. S3). Informed consent was obtained from every individual recruited, and this study was approved by the Eijkman Institute Research Ethics Commissions (EIREC No. 23/2007). Additional S gene sequences from 48 HBV/C isolates from Asia and 74 from the Papua-Pacific [10 from PNG, 20 from Vanuatu, 20 from Tonga, 20 from Fiji, and 4 from Kiribati islands] were downloaded from GenBank. Thus, together with the 87 newly generated and the initial 62 HBV/C isolates, we examined 271 sequences for S gene analysis (Table 2).

HBV DNA was extracted from 140 µL serum sample using QIAamp® viral DNA Mini Kit (Qiagen Inc., Chatsworth, CA) according to the manufacturer's instruction. PCR

amplification of the S region (226 bp) was carried out following a nested strategy using two oligonucleotide primer pairs: S2-1 (5'-CAAGGTATGTTGCCCGTTTG-3', nt 455-474) and S1-2 (5'-CGAACCACTGAACAAATGGC-3', nt 704-685) for the first round; S088 (5'-TGTTGCCCGTTTGTCCTCTA-3', nt 462-471) and S2-2 (5'-GGCACTAGTAAACTGAGCCA-3', nt 687-668) for the second round (Okamoto *et al*., 1988, Okamoto & Nishizawa,1992). Denaturizing, annealing and extension were carried out at 94°C for 30 s, 55°C for 30 s and 72°C for 1 min, respectively, for both rounds of PCR (35 cycles for the first and 25 for the second round). Amplification products were directly sequenced using Big Dye Terminator Reaction kits with ABI 3130 XL genetic analyzer (ABI Perkin Elmer, Norwalk, CT, USA).

HBsAg subtype determination of HBV/C strains

Deduced amino acid sequences of the S gene from the 271 HBV/C isolates were aligned using BioEdit package version 7.0 software. Amino acid variations that determine HBsAg subtypes (*adw*, *adr*, *ayw*, and *ayr*) were identified based on the common antigenic determinant *'a'* at amino acids 124-147 (Fig. S3), and two pairs of mutually exclusive determinants, d/y and w/r, at amino acids 122 and 160, respectively (Okamoto *et al*., 1987). Further specification into nine subtypes (*ayw*1, *ayw*2, *ayw*3, *ayw*4, *ayr, adw*2, *adw*3, *adw*4, *adrq*+ and *adrq*-) based on previous reports was also accomplished (Couroucé-Pauty *et al*., 1978; Norder *et al*., 1992).

Analysis of HBV/C surface and core immunoepitopes

Based on the sequence availability, analysis of surface immune epitopes for known recognition sites encompassing residues 20-180 of HBsAg was accomplished in 184 of the 271 HBV/C isolates, while the shorter sequence of 87 isolates generated in this study allowed only for B cell epitope analysis within residues 124 -148. Analysis of HBV/C core immunoepitopes was done for 147 isolates including 37 core sequences available from the Pacific, but not for the 87 sequences from this study due to insufficient volume of the repository specimens.The analysis was performed by comparing the CTL recognition sites, as well as T helper and B cell immunoepitopes of HBV/C isolates from East and Southeast Asia (Japan, Korea, China, Hongkong, Vietnam, Myanmar, Thailand, Malaysia, and Indonesia), and Papua-Pacific region (Papua New Guinea, Polynesia, New Caledonia, as well as Vanuatu, Fiji, and Tonga islands) with AF 473543 from China used as the reference [Fig.S1 and S2].

Acknowledgements

This study was supported by a national grant from the Ministry of Research and Technology, Indonesia. Reagents were partly contributed from PT Schering-Plough, Indonesia. The authors cordially thank Prof. Herawati Sudoyo, MD, Ph.D and team for consultation, ethnic identification and donation of samples. They would also like to express their gratitude to Prof. Julius and colleagues from Department of Internal Medicine, Faculty of Medicine, Andalas University, Padang, Indonesia; dr. Yuyun Soedarmono, Ph.D from the Indonesian Red Cross; dr. Zein Patiiha from Department of Medicine, Ternate General Hospital, Ternate, Indonesia, for the donation of samples; and to all related institutions for their support and collaboration.

References

1. **Bellwood, P. (1997).** Recent Indo-Malaysian prehistory: according to the languages. In *Prehistory of the Indo-Malaysian archipelago*, pp 96-127. edited by Bellwood, P., Fox, J.J., Tryon, D.: University of Hawai'i Press, Honolulu.
2. **Bertoletti, A., Costanzo, A., Chisari, F.V., Levrero, M., Artini, M., Sette, A., Penna, A., Giuberti, T., Fiaccadori, F. & Ferrari, C. (1994).** Cytotoxic T lymphocyte response to a wild type hepatitis B virus epitope in patients chronically infected by variant viruses carrying substitutions within the epitope. *J Exp Med* **180**:933-43.
3. **Bertoletti, A., Scott, S., Chesnut, R., Sette, A., Falco, M., Ferrara, G.B., Penna, A., Boni, C., Fiaccadori, F. & other authors. (1997).**Molecular Features of the Hepatitis B Virus Nucleocapsid T-Cell Epitope 18-27: Interaction With HLA and T-Cell Receptor. *Hepatology* **26**,1027-1034.
4. **Cao, G.W. (2009).** Clinical relevance and public health significance of hepatitis B virus genomic variations. *World J Gastroenterol* **15,** 5761-5769.
5. **Carman, W.F., Boner, W., Fattovich, G., Colman, K., Dornan, E.S., Thursz, M. & Hadziyannis, S. (1997).** Hepatitis B virus core protein mutations are concentrated in B cell epitopes in progressive liver disease and in T helper cell epitopes during clinical remission. *J Infect Dis* **175,** 1093–100.
6. **Carman, W.F., Jazayeri, M., Basuni, A., Thomas, H.C. & Karayiannis, P. (2005).** Hepatitis B surface antigen (HBsAg) variants. *In: Viral hepatitis. 3rd.ed.,* pp 225-241. Edited by: Thomas, H.C., Lemon, S. & Zuckerman, A.: Blackwell Publishing Ltd, Oxford, UK.
7. **Cavinta, L., Sun, J., May, A., Yin, J., von Meltzer, M., Radtke, M., Barzaga, N.G., Cao, G. & Schaefer, S. (2009).** A new isolate of hepatitis B virus from the Philippines possibly representing a new subgenotype C6. *J Med Virol* **81**, 983–987.
8. **Chan, H.L.Y., Tsui, S.K.W., Tse, C.H., Ng, E.Y.T., Au, T.C.C., Yuen, L., Bartholomeusz, A., Leung, K.S., Lee, K.H. & other authors. (2005).**

Epidemiological and virological characteristics of 2 subgroups of hepatitis B virus genotype C. *J Infect Dis* **191**, 2022-2032.

9. **Chisari. (2000).** Viruses, immunity, and cancer: lessons from hepatitis B. Rous-Whipple Award Lecture. *Am J Pathol* **156**, 1118-1132.

10. **Couroucé-Pauty, A.M., Lemaire, J.M. & Roux, J.F. (1978).** New hepatitis B surface antigen subtypes inside the *ad* category. *Vox Sang* **35**, 304-308.

11. **Deka, R., Su, B. & Jin, L. (2001).** The genetic trail from Southeast Asia to the Pacific. In: Genetic, linguistic, and archeological perspectives on human diversity in Southeast Asia, pp 135-145. Edited by: Jin, L., Seielstad, M. & Xiao, C.: World Scientific Publisging Co, Pte, Ltd, Singapore.

12. **Ehata, T., Omata, M., Yokosuka, O., Hosoda, K. & Ohto, M. (1991).** Amino acid residues of core region of hepatitis B virus. Asymptomatic carriers versus patients with liver disease. *J Gastroenterol Hepatol* **6**, 292-6.

13. **Ehata, T., Omata, M., Yokosuka, O., Hosoda, K. & Ohto, M. (1992).** Variations in codons 84-101 in the core nucleotide sequence correlate with hepatocellular injury in chronic hepatitis B virus infection. *J Clin Invest* **89**, 332-8.

14. **Ferrari, C., Bertoletti, A., Penna, A., Cavalli, A., Valli, A., Missale, G., Pili, M., Fowler, P., Giuberti, T. & other authors. (1991).** Identification of immunodominant T cell epitopes of the hepatitis B virus nucleocapsid antigen. *J Clin Invest* **88**, 14-22.

15. **Friedlaender, J.S., Friedlaender, F.R., Reed, F.A., Kidd, K.K., Kidd, J.R., Geoffrey, K., Chambers, G.K., Lea, R. A., Loo, J.H., Koki, G., Hodgson, J.A., Merriwether, M.A., Weber, L.L. (2008).** The genetic structure of Pacific Islanders. *PLoS Genet* **4**, e19. doi: 10.1371/journal. pgen.0040019

16. **Huy, T.T., Ushijima, H., Quang, V.X., Luengrojanakul, P., Kikuchi, K., Sata, T. & Abe, K. (2004).** Genotype C of hepatitis B virus can be classified into at least two subgroups. *J Gen Virol* **85**, 283-292.

17. **Huy, T.T.T., Ngoc, T.T. & Abe, K. (2008).** New complex recombinant genotype of Hepatitis B Virus identified in Vietnam. *J Virol* **82**, 5657-5663.

18. **Jazayeri, M.S., Basuni, A.A., Cooksley, G., Locarnini, S. & Carman, W.F. (2004a).** Hepatitis B virus genotypes, core gene variability and ethnicity in the Pacific region. *J Hepatol* **41**, 139-146.

19. **Jazayeri, M.S., Basuni, A. A., Sran, N., Gish, R., Cooksley, G., Locarnini, S., & Carman, W.F. (2004b).** HBV core sequence: definition of genotype-specific variability and correlation with geographical origin. *J Viral Hepat* **11**, 488–501

20. **Jazayeri, M.S. & Carman, W.F. (2005).** Virus escape CTL or B cell epitopes ? *Hepat Monthly* **5**, 133-136.

21. **Kao, J.H., Chen, P.J., Lai, M.Y. & Chen, D.H. (2000).** Hepatitis B genotypes correlate with clinical outcome in patients with chronic hepatitis B. *Gastroenterol* **118,** 554-559.

22. **Kao, J.H. (2011).** Molecular epidemiology of hepatitis B virus. *Korean J Intern Med* **26**, 255-261.

23. **Kim, H.J., Lee, D.H., Gwak, G.Y., Choi, M.S., Lee, J.H., Koh, K.C., Paik, S.W. & Yoo, B.C. (2007).** Analysis of the core gene of hepatitis B virus in Korean patients.

Liver Inter **27,** 633-638.

24. **Kurbanov, F., Tanaka, Y. & Mizokami, M. (2010).** Geographical and genetic diversity of the human hepatitis B virus. *Hepatology Research* **40**, 14-30.

25. **Lavanchy, D.** (2004). Hepatitis B virus epidemiology, disease burden, treatment, and current and emerging prevention and control measures. *J Viral Hepat* **11**: 97–107.

26. **Lusida, M.I., Nugrahaputra, V.E., Soetjipto, R., Handajani, M., Nagano-Fujii, M., Sasayama, T., Utsumi, T. & Hotta, H. (2008).** Novel subgenotypes of hepatitis B virus genotypes C and D in Papua, Indonesia. *J Clin Microbiol* **46**, 2160-2166.

27. **Marzuki, S., Sudoyo, H., Suryadi, H., Setianingsih, I., & Pramoonjago, P. (2003).** Human genome diversity and disease on the island Southeast Asia. In *Tropical diseases: from molecule to bedside*, pp 3-18, Adv Exp Med Biol 531. Edited by Marzuki, S., Verhoef, J., Snippe, H. Kluwer Academic/Plenum Publishers, New York.

28. **McMahon, G., Ehrlich, P.H, Moustafa, Z.A., McCarthy, L.A., Dottavio, D., Tolpin, M.D., Nadler, P.I. & Ostberg, L. (1992).** Genetic alterations in the gene encoding the major HBsAg: DNA and immunological analysis of recurrent HBsAg derived from monoclonal antibody-treated liver transplant patients. *Hepatology* **15**,757–766.

29. **Min, W.P., Kamikawaji, N., Mineta, M., Tana, T., Kashiwagi, S. & Sasazuki, T. (1996).** Identification of an epitope for T-cells correlated with antibody response to Hepatitis B surface Antigen in vaccinated humans. *Human Immunol* **46**, 93-99.

30. **Miyakawa, Y. & Mizokami, M. (2003).** Classifying hepatitis B virus genotypes. *Intervirology* **46**, 329-338.

31. **Mohamadkhani, A., Jazii, F.R., Poustchi, H., Nouraein, O., Abbasi, S., Sotoudeh, M. & Montazeri, G. (2009).** The role of mutations in core protein of hepatitis B virus in liver fibrosis. *Virology* **2009**, 209-217.

32. **Mulyanto, Tsuda, F., Karossi, A.T., Soewignjo, S., Roestamsjah, Sumarsidi, D., Trisnamurti, R.H., Sumardi, Surayah, & other authors. (1997).** Distribution of the hepatitis B surface antigen subtypes in Indonesia: implication of ethnic heterogeneity and infection control measures. *Arch Virol* **142**, 2121-2129.

33. **Mulyanto, Depamede, S.N., Surayah, K., Tsuda, F., Ichiyama, K., Takahashi, M. & Okamoto, H. (2009).** A nationwide molecular epidemiological study on hepatitis B virus in Indonesia: identification of two novel subgenotypes, B8 and C7. *Arch Virol* **154**,1047-1059.

34. **Mulyanto, Depamede, S.N., Surayah, K., Tjahyono, A.A.H., Jirintai, Nagashima, S., Takahashi, M. & Okamoto, H. (2010).** Identification and characterization of novel hepatitis B virus subgenotype C10 in Nusa Tenggara, Indonesia. *Arch Virol* **155**, 705-715.

35. **Mulyanto, Depamede, S.N., Wahyono, A., Jirintai, Nagashima, S., Takahashi, M. & Okamoto, H. (2011).** Analysis of the full-length genomes of novel hepatitis B virus subgenotypes C11 and C12 in Papua, Indonesia. *J Med Virol* **83**, 54-64.

36. **Mulyanto, Pancawardani, P., Depamede, S.N., Wahyono, A., Jirintai, Nagashima, S., Takahashi, M., Nishizawa, T. & Okamoto, H. (2012).** Identification of four novel subgenotypes (C13-C16) and two inter-genotypic recombinants (C12/G and C13/B3) of hepatitis B virus in Papua Province, Indonesia. *Virus Res* **163**,129-140.

37. **Nayersina, R., Fowler, P., Guilhot, S., Missale, G., Cerny, A., Schlicht, H.J., Vitiello, A., Chesnut, R., Person, J.L. & other authors. (1993).** HLA A2 restricted cytotoxic T lymphocyte responses to multiple hepatitis B surface antigen epitopes during hepatitis B virus infection. *J Immunol* **150**, 4659-4671.

38. **Nurainy, N., Muljono, D.H., Sudoyo, H. & Marzuki, S. (2008).** Genetic study of hepatitis B virus in Indonesia reveals a new subgenotype of genotype B in east Nusa Tenggara. *Arch Virol* **153**, 1057-1065.

39. **Norder, H., Couroucé, A.M. & Magnius, L.O. (1992).** Molecular basis of hepatitis B virus serotype variations within the four major subtypes. *J Gen Virol* **73**, 3141-3145.

40. **Norder, H., Couroucé, A.M. & Magnius, L.O. (1994).** Complete genome, phylogenetic relatedness, and structural proteins of six strains of the hepatitis B virus, four of which represent two new genotypes. *Virology* **198**, 489-503.

41. **Norder, H., Couroucé, A.M., Coursaget, P., Echevarria, J.M., Lee, S.D., Mushahwar, I.K., Robertson, B.H., Locarnini, S. & Magnius, L.O. (2004).** Genetic diversity of hepatitis B virus strains derived worldwide: Genotypes, Subgenotypes, and HBsAg subtypes. *Intervirology* **47**, 289-309.

42. **Okamoto, H., Imai, M., Tsuda, F., Tanaka, T., Miyakawa, Y. & Mayumi, M. (1987).** Point mutation in the S gene of Hepatitis B Virus for a d/y or w/r subtypic change in two blood donors carrying a surface antigen of compound subtype adyr or adwr. *J Virol* **61**, 3030-3034.

43. **Okamoto, H., Tsuda, F., Sakugawa, H., Sastrosoewignjo, R.I., Imai, M., Miyakawa, Y. & Mizokami, M. (1988)**. Typing hepatitis B virus by homology in nucleotide sequence: comparison of surface antigen subtypes. *J Gen Virol* **69**, 2575-2583.

44. **Okamoto, H. & Nishizawa, T. (1992).** Non-B non-G hepatitis virus gene, polynucleotide, polypeptide, virion, method for separating virion, and method for detecting virus. http://www.freepatentsonline.com/ EP1010759.html (with permission).

45. **Orito, E., Mizokami, M., Sakugawa, H., Michitaka, K., Ishikawa, K., Ichida, T., Okanoue, T., Yotsuyanagi, H. & Iino, S. (2001).** A Case-control study for clinical and molecular biological differences between Hepatitis B Viruses of genotypes B and C. *Hepatology* **33**, 218-223.

46. **Pairan, A. & Bruss, V. (2009).** Functional surfaces of the hepatitis B virus capsid. *J Virol* **83**, 11616-11623.

47. **Sakamoto, T., Tanaka, Y., Orito, E., Co, J., Clavio, Sugauchi, F., Ozasa, A., Quino, A., Ueda, R. & other authors (2006).** Novel subtypes (subgenotypes) of hepatitis B virus genotypes B and C among chronic liver disease patients in the Philippines. *J Gen Virol* **87**, 1873-1882.

48. **Sastrosoewignjo, R.I., Sandjaya, B. & Okamoto, H. (1991).** Molecular epidemiology of hepatitis B virus in Indonesia. *J GastroenterolHepatol* **6**, 491-498.

49. **Sugauchi, F., Mizokami, M., Orito, E., Ohno, T., Kato, H., Suzuki, S., Kimura, T., Ueda, R., Butterworth, L.A. & other authors. (2001).** A novel variant genotype C hepatitis B virus identified in isolates from Australian Aborigines: Complete genome sequence and phylogenetic relatedness. *J Gen Virol* **82**, 883-892.

50. **Tai, P.C., Banik, D., Lin, G.I., Pai, S., Pai, K., Lin, M.L., Yuoh, G.B.O., Che, S., Hsu, S.H. & other authors. (1997).** Novel and Frequent Mutations of Hepatitis B Virus Coincide with a Major Histocompatibility Complex Class I-Restricted T-Cell Epitope of the Surface Antigen. *J Virol* **71**, 4852-4856.

51. **Tan, A.T., Loggi, E., Boni, C., Chia, A., Gehring, A.J., Sastry, K.S.R., Goh, V., Fisicaro, P., Andreone, P. & other authors (2008).** Host Ethnicity and Virus Genotype Shape the Hepatitis B Virus-Specific T-Cell Repertoire. *J Virol* **82**, 10986-10997.

52. **Tatematsu, K., Tanaka, Y., Kurbanov, F., Sugauchi, F., Mano, S., Maeshiro, T., Nakayoshi, T., Wakuta, M., Miyakawa, Y. & other authors (2009).** A genetic variant of Hepatitis B Virus divergent from known human and ape genotype isolated from a Japanese patient and provisionally assigned to new genotype J. *J Virol***83**, 10538-10547.

53. **Thedja, M.D., Roni, M., Harahap, A.R., Siregar, N.C., Ie, S.I. & Muljono, D.H. (2010).** Occult hepatitis B in blood donors in Indonesia: altered antigenicity of the hepatitis B virus surface protein.*Hepatol Int* **4**, 608-614.

54. **Thedja, M.D., Muljono, D.H., Nurainy, N., Sukowati, C.H.C., Verhoef, J. & Marzuki, S. (2011).** Ethnogeographical structure of hepatitis B virus genotype distribution in Indonesia and discovery of a new subgenotype, B9. *Arch Virol* **156**, 855-868.

55. **Thursz, M., Yee, L. & Khakoo, S. (2011).** Understanding the host genetics of chronic hepatitis B and C. *Sem Liver Dis* **31**, 115-127.

56. **Tiollais, P., Pourcel, C. & Dejean, A. (1985).** The hepatitis B virus. *Nature* **317**,489–495.

57. **Tryon, D. (1995).** Proto-Austronesian and the major Austronesian subgroups. In *The Austronesians: historical and comparative perspectives*, pp 17-35. Edited by Bellwood, P., Fox, J.J. & Tryon, D: The Australian National University, Canberra.

58. World Health Organization (WHO). (2012). Hepatitis B Fact Sheet No. 204.http://www.who.int/mediacentre/factsheets/fs204/en/.

59. **Yang, H.I., Yeh, S.H., Chen, P.J., Iloeje, H.I., Jen, C.L., Su. J., Wang, L.Y., Lu, S.N., You, S.L., and other authors (2008).** Associations Between Hepatitis B Virus Genotype and Mutants and the Risk of Hepatocellular Carcinoma. *J Natl Cancer Inst* **100**, 1134 – 1143.

60. **Zeng, G., Wang, Z., Wen, S., Jiang, J., Wang, L., Cheng, J., Tan, D., Xiao, F., Ma, S. & other authors. (2005).** Geographic distribution, virologic and clinical characteristics of hepatitis B virus genotype in China. *J Viral Hepat* **12**, 609-617.

Table 1. Mean percentage nucleotide divergence of the complete genome between HBV/C subgenotypes.

	C1 (27)	C2 (36)	C3 (2)	C4 (2)	C5 (8)	C6 (14)	C8 (4)	C11 (2)	C12 (4)	C13 (3)	C14 (2)
C1	**2.68 ± 0.65**										
C2	4.37 ± 0.52	**2.52 ± 0.62**									
C3	4.86 ± 0.48	4.41 ± 0.34	**2.71 ± 1.91**								
C4	7.16 ± 0.41	6.36 ± 0.35	6.01 ± 0.24	**0.91 ± 0.64**							
C5	5.65 ± 0.39	5.27 ± 0.40	5.92 ± 0.32	7.71 ± 0.30	**1.55 ± 0.48**						
C6	5.78 ± 0.44	4.66 ± 0.33	4.49 ± 0.27	6.57 ± 0.25	6.36 ± 0.28	**2.53 ± 0.94**					
C8	5.30 ± 0.38	4.28 ± 0.28	4.71 ± 0.24	6.26 ± 0.22	5.59 ± 0.23	4.94 ± 0.23	**0.77 ± 0.38**				
C11	5.64 ± 0.43	4.71 ± 0.24	4.49 ± 0.20	6.19 ± 0.22	6.44 ± 0.19	4.02 ± 0.17	4.95 ± 0.14	**0.19 ± 0.13**			
C12	6.05 ± 0.51	5.07 ± 0.51	5.03 ± 0.50	6.69 ± 0.42	6.62 ± 0.39	4.98 ± 0.56	5.28 ± 0.41	4.91 ± 0.56	**3.25 ± 1.58**		
C13	6.19 ± 0.41	5.64 ± 0.36	5.57 ± 0.28	7.44 ± 0.33	7.29 ± 0.29	5.01 ± 0.41	5.75 ± 0.29	4.72 ±0.21	5.92 ± 0.48	**2.26 ± 0.46**	
C14	4.61 ± 0.42	3.91 ± 0.31	4.83 ± 0.17	6.77 ± 0.16	5.78 ± 0.29	5.03 ± 0.11	4.82 ± 0.09	5.25 ± 0.06	5.36 ± 0.53	5.64 ± 0.25	**0.75 ± 0.53**

#The total number of HBV/C isolates examined of each subgenotype is shown in bracket. Other existing subgenotypes (C7, C9, C10, C15, and C16) were not included in the genetic distance calculation since only single isolate was available for each subgenotype. Intragenotype divergences are shown in bold.

Table 2. Characterization of subgenotypes and subtypes of 271 HBV/C isolates according to their country/geographical origins involving East/Southeast Asia and Papua-Pacific.

Origins	N	Subtype	n	Genotype/subgenotype																
				C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C
China	16	*adrq+*	15	2	13															
		adw2	1		1															
Japan	21	*adrq+*	20		20															
		adr_indet (A159/A177)	1		1															
Korea	3	*adrq+*	2		2															
		adr_indet (A159/A177)	1		1															
Myanmar	8	*adrq+*	7	7																
		adw2	1	1																
Thailand	9	*adrq+*	7	7																
		adr_indet (V159/V177)	1	1																
		ayr	1	1																
Vietnam	5	*adrq+*	5	5																
Malaysia	1	*adrq+*	1	1																
Philippine	5	*adw2*	4					4												
		adrq+	1							1										
Indonesia	124#	*adrq+*	78	38	19				9		4		1		4		2	1		
		adw2	17	5	3			9												
		ayr	14	12										2						
		ayw1	2	1	1															
		adr_indet (A159/A177)	13						9							3			1	
Timor Leste	1	*adrq+*	1									1								
Australia	2	*ayw3*	2				2													
Polynesia	1	*adrq-*	1			1														
N. Caledonia	1	*adrq-*	1			1														
PNG	10	*adrq+*	8																	8
		adr_indet (V159/V177)	2																	2
Vanuatu	20	*adrq-*	16																	16
		adrq+	1																	1
		adr_indet (V159/V177)	3																	3
Fiji	20	*adrq-*	13																	13
		adr_indet (V159/V177)	6																	6
		ayw1	1																	1
Tonga	20	*adrq-*	19																	19
		adr_indet (V159/V177)	1																	1
Kiribati	4	*adrq-*	3																	3
		adrq+	1																	1
TOTAL			**271**	**81**	**61**	**2**	**2**	**13**	**18**	**1**	**4**	**1**	**1**	**2**	**4**	**3**	**2**	**1**	**1**	**74**

i# including 37 published complete genome sequences and 87 newly generated in this study

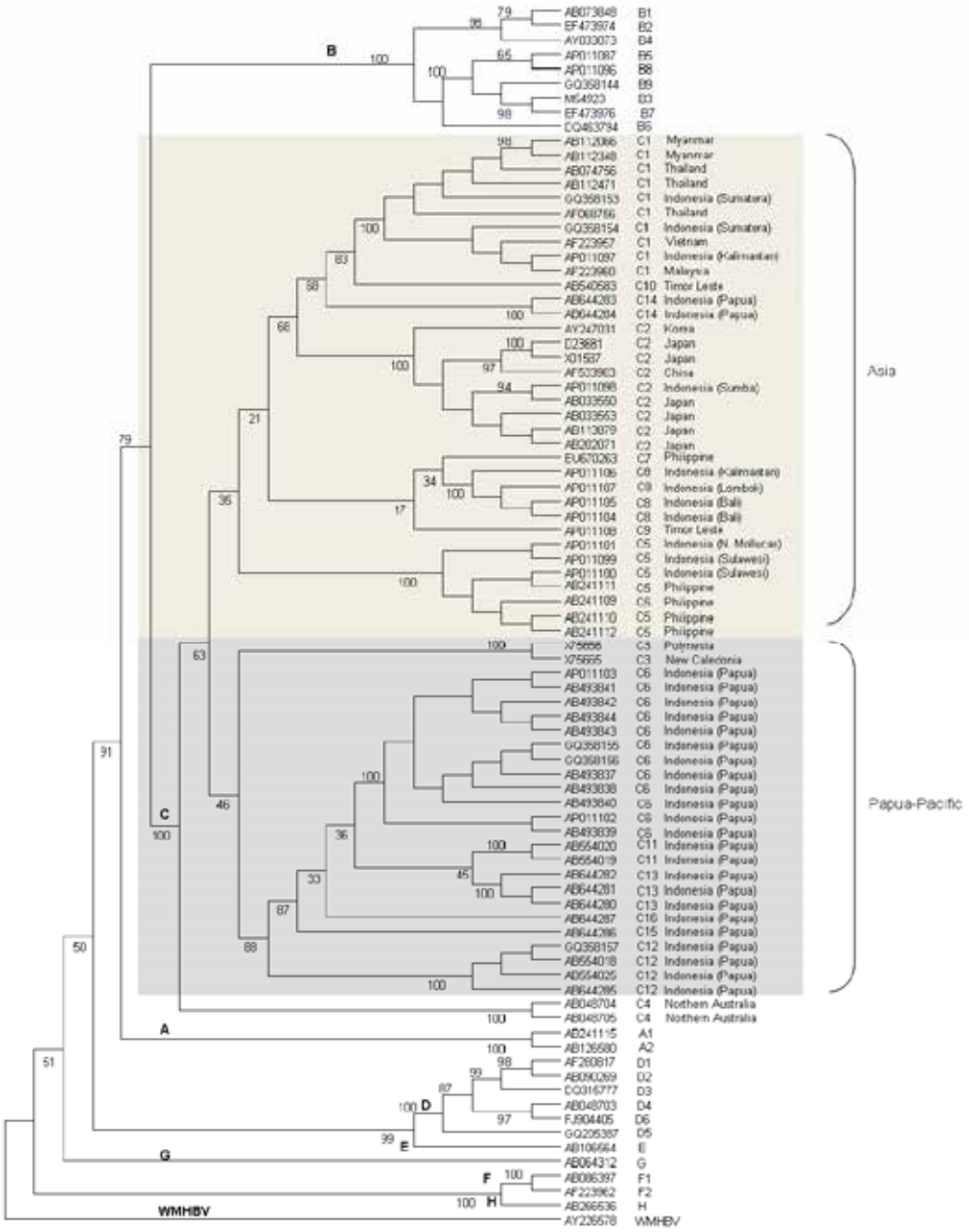
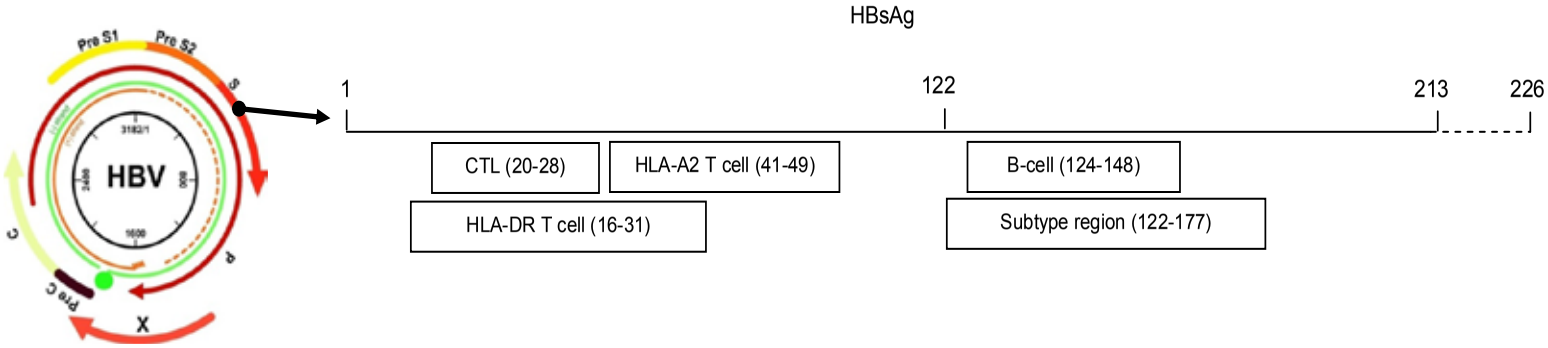
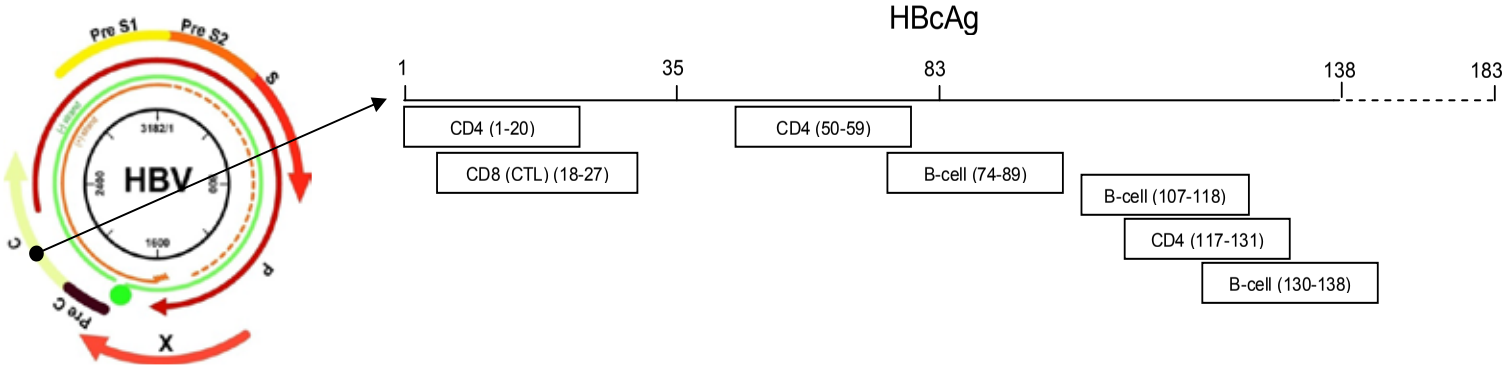


Figure 1. Phylogenetic tree of the HBV/C complete genome sequences of isolates from different countries in East and Southeast Asia, and Papua-Pacific. Isolates from various subgenotypes (C1-C16) are clearly grouped into two major clusters and consistent with their geographical origins. Seven HBV/C subgenotypes (C1, C2, C5, C7, C8, C9, and C10) from East and Southeast Asia, and one (C14) from Papua (*light highlight*) were well-separated from those seven subgenotypes (C3, C6, C11, C12, C13, C15, and C16) from Papua-Pacific (*dark highlight*). The diversification of the Asian type from the Papua-Pacific type started from Papua of Indonesia to the east. The other subgenotype, HBV/C4, was distanced from other subgenotypes. In this analysis, one strain (GQ358157) from Papua reported as C6 in our previous study (Thedja *et al*., 2011) grouped into C12. We redefine this strain as a member of HBV/C12.



Subgenotype	Country/Origin	n	Amino acid position															
			18	20	24	27	28	40	44	47	122	126	127	134	159	160	177	213
C1	Asia	28	G	F	R	T	I	N	G	T	K	I	P	F	A	R	V	L
C2	Asia	39	G	F	R	T	I	N	G	T	K	I	P	F	A	R	V	L
C5	The Philippine (SEA)	7	G	F	K	T	I	N	G	T	K	T	P	F	A	K	A/V	L/I/M
C7	The Philippine (SEA)	1	G	F	K	T	I	N	G	M	K	I	P	F	A	R	V	I
C8	Indonesia (Kupang)	4	G	F	K	T	I	N	G	V	K	I	P	F	A	R	V	L/I
C9	East Timor (Dili)	1	G	F	K	T	I	N	G	V	K	I	P	F	A	R	V	L
C10	Indonesia (East N. Tenggara)	1	G	F	R	T	I	N	G	T	K	I	P	F	A	R	V	L
C4	Northern Australia	2	G	F	R	T	I	N	G	G/A	R	I	T	F	A	K	V	L/I
C14	Indonesia (Papua)	2	G	F	R	T	I	N	E	K/T	K	I	P	F	A	R	V	L
C3	Polynesia, N. Caledonia	2	V	F	R/K	T	I	N	E/G	R/V	K	I	P	F	V	R	A	I
C6	Indonesia (Papua)	12	V	F	R	T	I	N	E	M	K	I	P	F	A	R	A/V	I
C11	Indonesia (Papua)	2	V	F	R	T	I	N	E	M	R	I	P	F	A	R	A	I
C12	Indonesia (Papua)	4	V	F	R	T	I	N	E	A	K	I	P	F	A	R	V	I
C13	Indonesia (Papua)	3	V	F	R	T	I	N	E	M	K	I	P	F	A	R	A	I
C15	Indonesia (Papua)	1	V	F	K	T	I	N	E	A	K	I	P	F	A	R	V	I
C16	Indonesia (Papua)	1	V	F	R	T	I	N	E	M	R	I	P	F	A	R	A	I
C PNG	Papua New Guinea	10	V	F	R	T	I	N	E	A	K	I	P	F	A/V	R	V	I
C sPacific	Vanuatu, Fiji, Tonga	64	V	F	R	T	I	N	E/G	V/R	K	I	P	F	V	R	A/V	I

Figure 2. The most frequent amino acid variations in B and T cell epitopes of HBcAg among 141 isolates from various HBV/C subgenotypes of Asia and Papua-Pacific regions. Among 15 amino acid positions examined within HBcAg immune epitopes, we identified I/V at position 59 as the essential variation that classified HBV/C subgenotypes into two major clusters, the Asian and the Papua-Pacific. HBV/C4 and C14 showed similar variation in most amino acids examined, with C4 and C14 having I59 and V59, respectively.



Subgenotype	Country/Origin	n	Amino acid position														
			12	27	35	40	59	67	74	83	87	91	97	107	118	130	138
C1	Asia	32	S	I/V	S	E	I	N	S	E	S/G	V	I/L	C	Y	P	P
C2	Asia	34	S	I/V	S	E	I	N	S	E	S	V	I/L	C	Y	P/T	P
C5	The Philippine (SEA)	8	S	I	S	E	I	N	S	E	S	V	I	C	Y	P	P
C7	The Philippine (SEA)	1	S	I	S	E	I	N	S	E	S	V	I	C	Y	P	P
C8	Indonesia (Kupang)	4	S	I	S	E	I	N	S	E	S	V	I	C	Y	P	P
C9	East Timor (Dili)	1	S	I	S	E	I	N	S	E	S	V	I	C	Y	P	P
C10	Indonesia (East N. Tenggara)	1	S	I	X	E	I	N	S	E	S	V	I	C	Y	Q	P
C4	Northern Australia	2	S	I	S	E	I	N	S	E	S	V	I	C	Y	I/L	P
C14	Indonesia (Papua)	2	S	I	S	E	V	N	S	E	S	V	I	C	Y	P/T	P
C3	Polynesia, N. Caledonia	2	S	I	S	E	V	N	S	E	S	I/V	I	C	Y	P	P
C6	Indonesia (Papua)	13	S	I	S	E	V	N	S	E	S	V	I	C	Y	P	P
C11	Indonesia (Papua)	2	S	I	S	E	V	N	S	E	S	V	I	C	Y	P/T	P
C12	Indonesia (Papua)	4	S	I	S	E	V	N	S	E	S	V	I	C	Y	P/L	P
C13	Indonesia (Papua)	2	S	I	S	E	V	N	S	E	S/G	V	L/I	C	T	Q/P	P
C15	Indonesia (Papua)	1	S	I	S	E	V	N	S	E	S	V	I	C	Y	P	P
C16	Indonesia (Papua)	1	S	I	S	E	V	N	S	E	S	V	I	C	Y	P	P
C sPacific	Vanuatu, Fiji, Tonga, Pacific	37	S	I	S	E	V	N	S	E	S	V/I	I	C	Y	P	P

Figure 3. The most frequent amino acid variation in B and T cell epitopes of HBsAg among 179 isolates from various HBV/C subgenotypes of Asia and Papua-Pacific regions. Analysis of HBV/C immune epitopes within HBsAg identified two patterns of amino acid variations corresponding to the geographical origins of the isolates. The first pattern of seven subgenotypes represented HBV/C subgenotypes from Asia (*light highlight*), while the second pattern of seven subgenotypes and unclustered subgenotypes from PNG and Pacific (*dark highlight*) was from Papua and Pacific. Two subgenotypes, HBV/C4 and C14, did not belong to either Asia or Papua-Pacific clusters. We also identified G18 and G44 as the critical points distinguishing the Asian from the Papua-Pacific type.

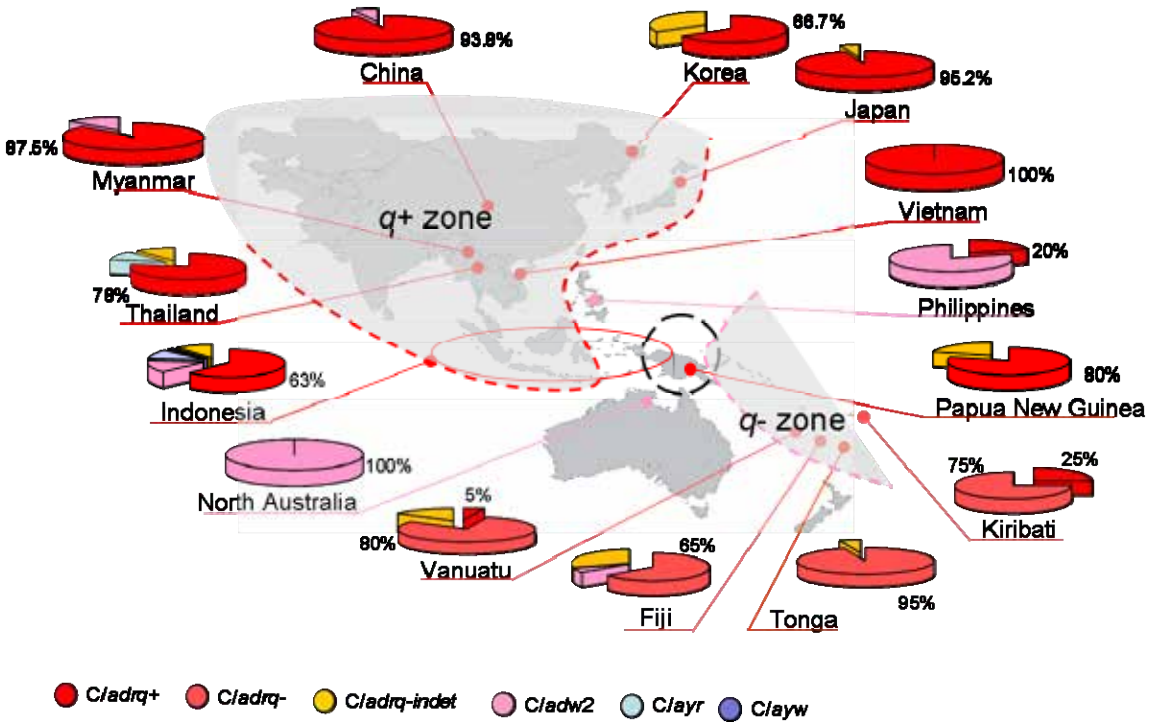
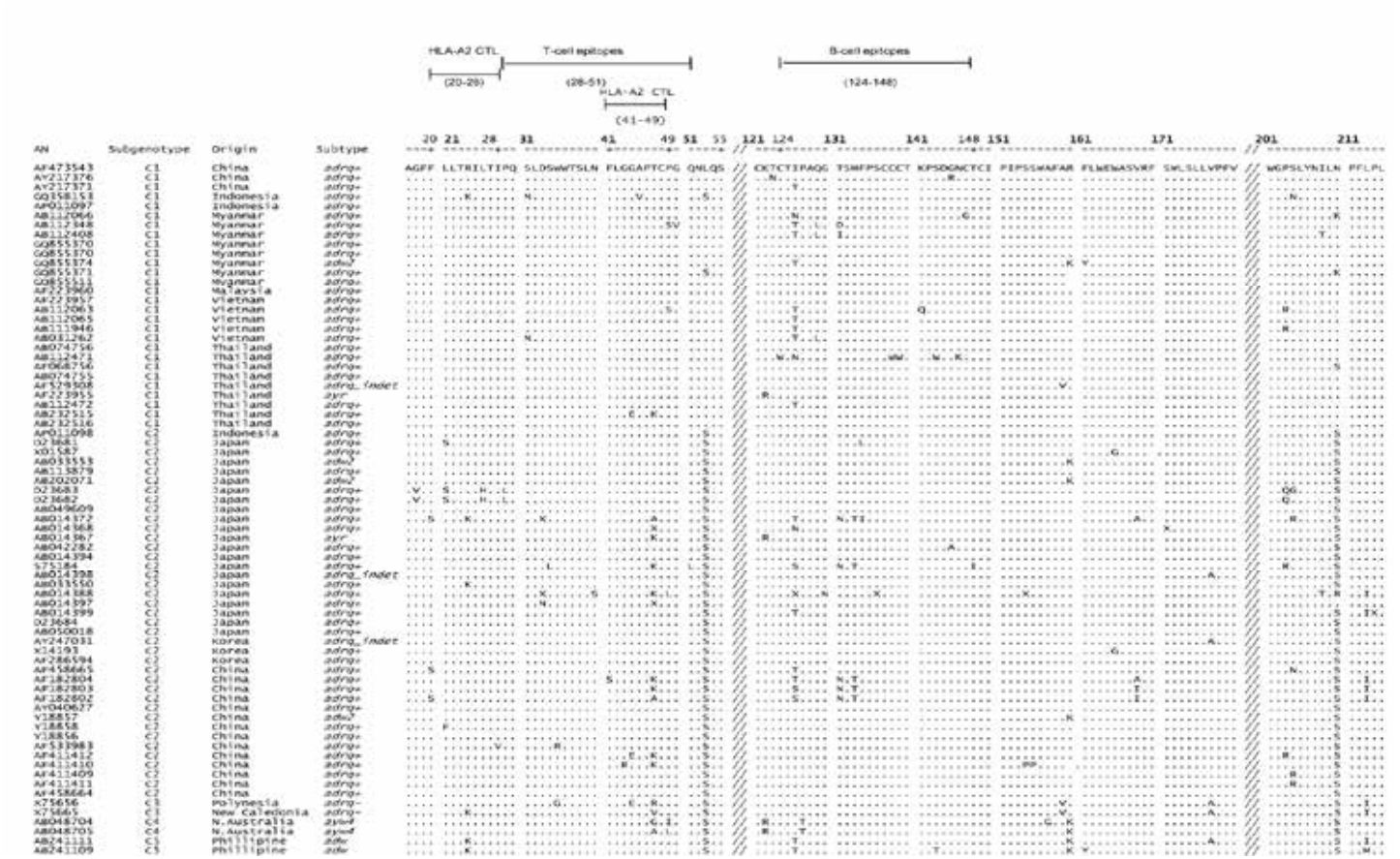


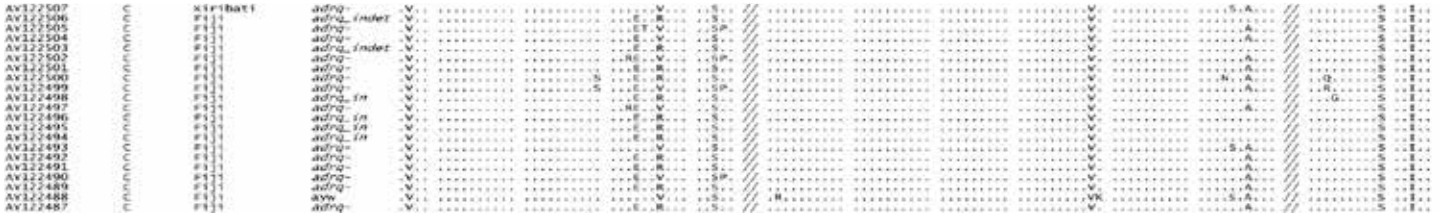
Figure 4. The distribution of HBV/C subtypes in the East and Southeast Asia and the Papua-Pacific. This study identified a west-to-east gradient in the distribution of HBsAg subtypes with *adrq+* (*red*) prominent in East-Southeast Asia and *adrq-* (*light pink*) in the Pacific region (Vanuatu, Fiji, Tonga, and Kiribati). Interestingly, together with *adrq+, adrq*-indeterminate A159/A177 and a new pattern of *adrq*-indeterminate V159/V177 identified in this study were found in Papua and PNG, respectively, suggesting that the molecular admixture of HBV/C, particularly for subtype evolution, occurred in Papua and PNG with both *adrq*-indeterminate forms as the transitional patterns.

Supplementary Figure 1



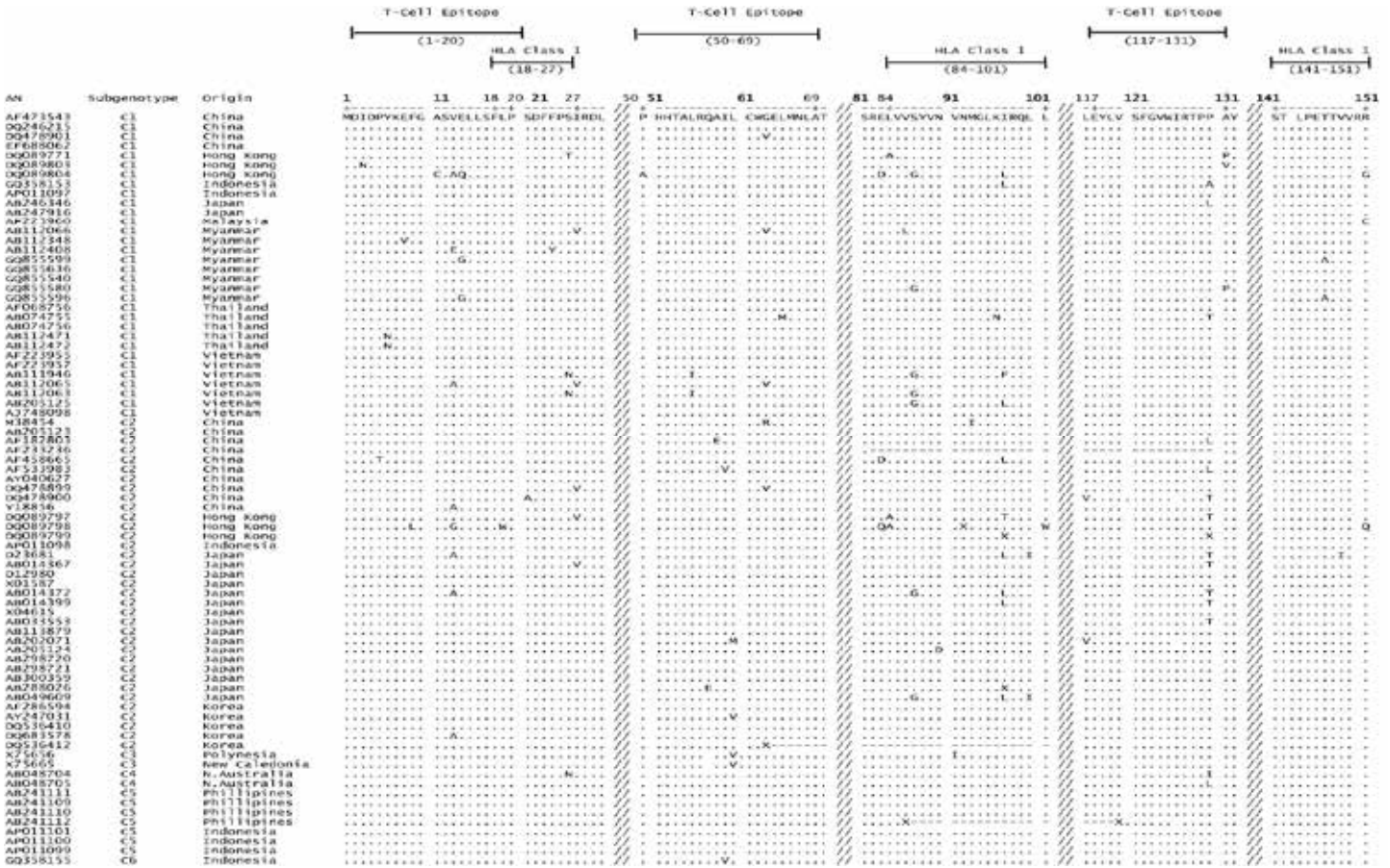
Supplementary Figure 1. continued

Supplementary Figure 1. continued

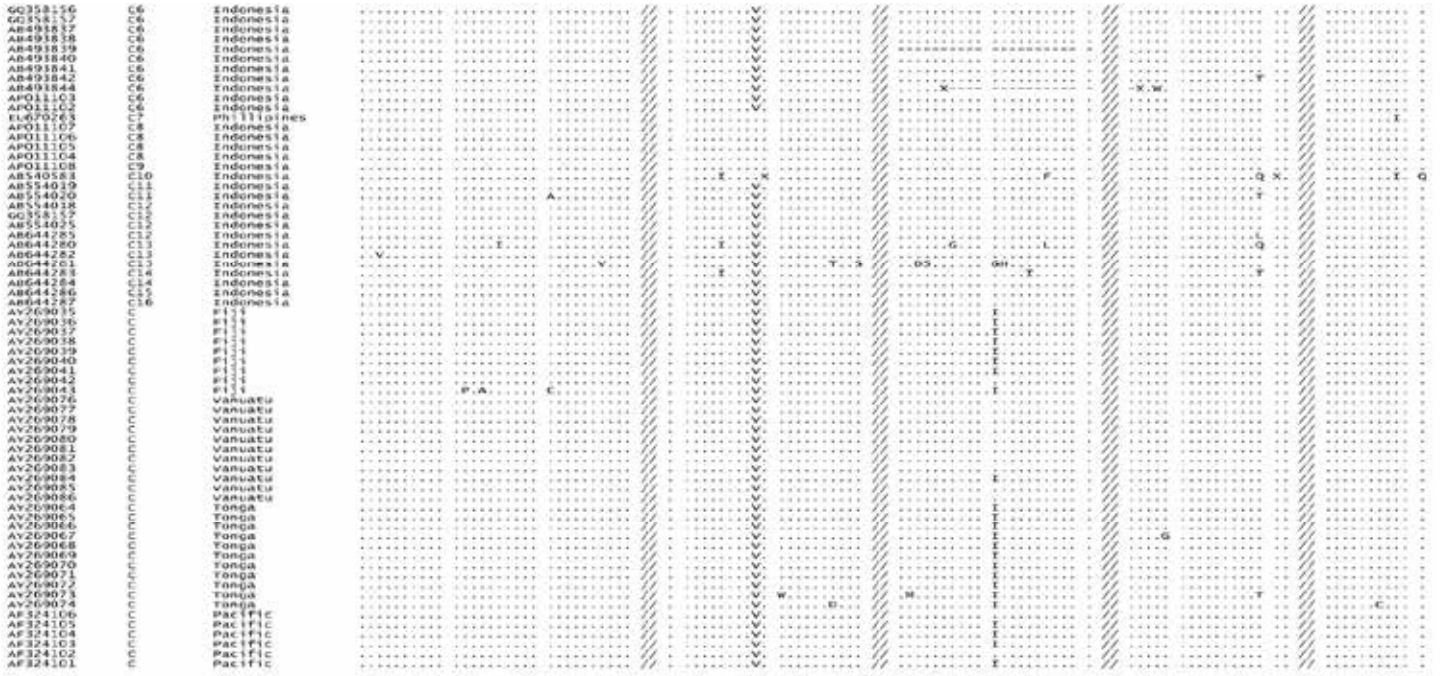


Amino acid sequence alignment of HBcAg of 141 HBV/C isolates. One hundred and forty-one isolates of various HBV/C subgenotypes were examined for immune epitopes within the core protein (HBcAg). The isolates were retrieved from GenBank following their origins from various geographical regions of the East and Southeast Asia and the Papua-Pacific. Dots indicate amino acids identical to those of AF473543 of China used as consensus. The types and location of immune epitopes in HBcAg are indicated. In general, HBV/C isolates from East and Southeast Asia have higher amino acid variation than those from Papua-Pacific. A single amino acid variation—I59V—markedly demonstrates the clustering of isolates from Asia and Papua-Pacific.

Supplementary Figure 2

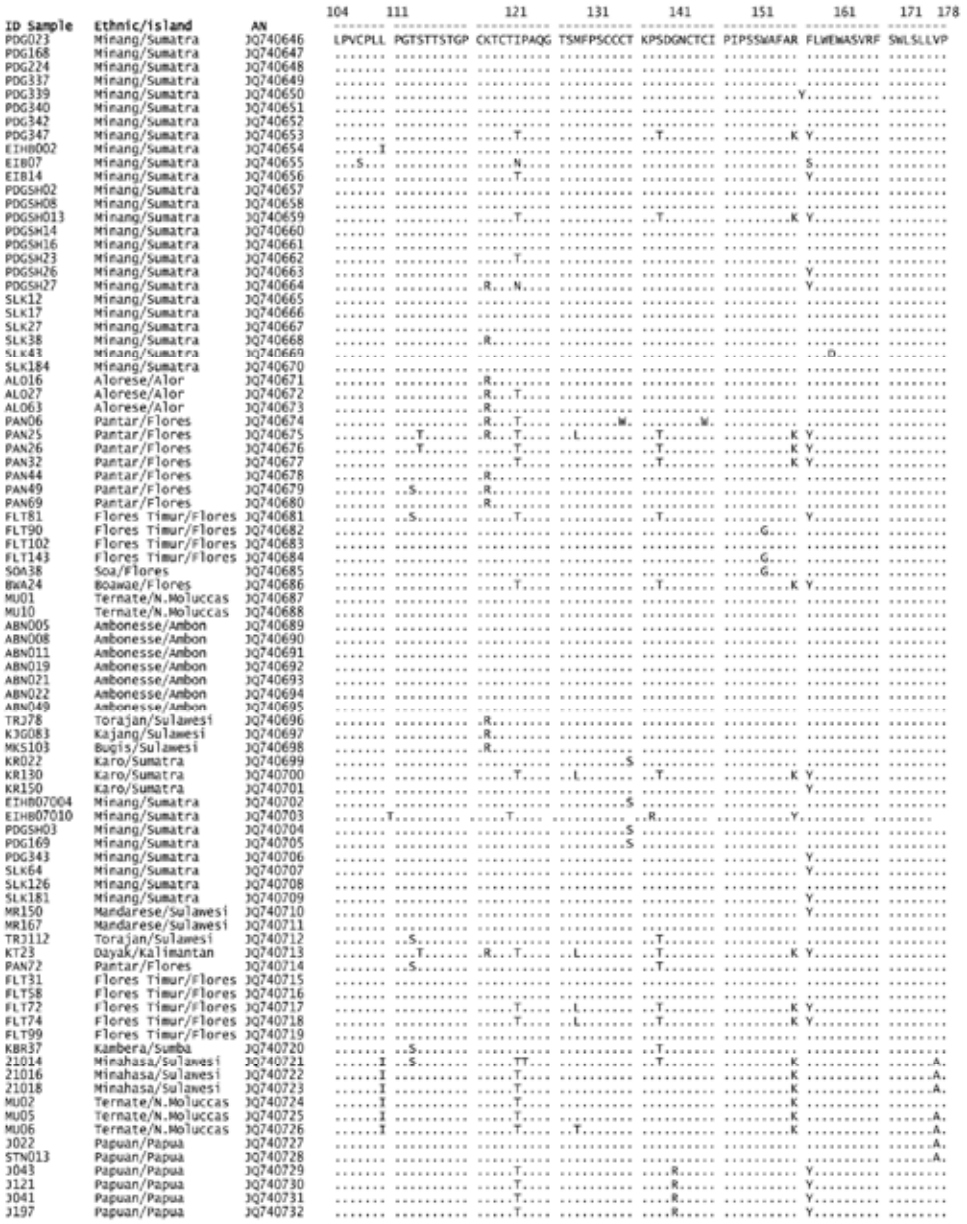


Supplementary Figure 2. continued



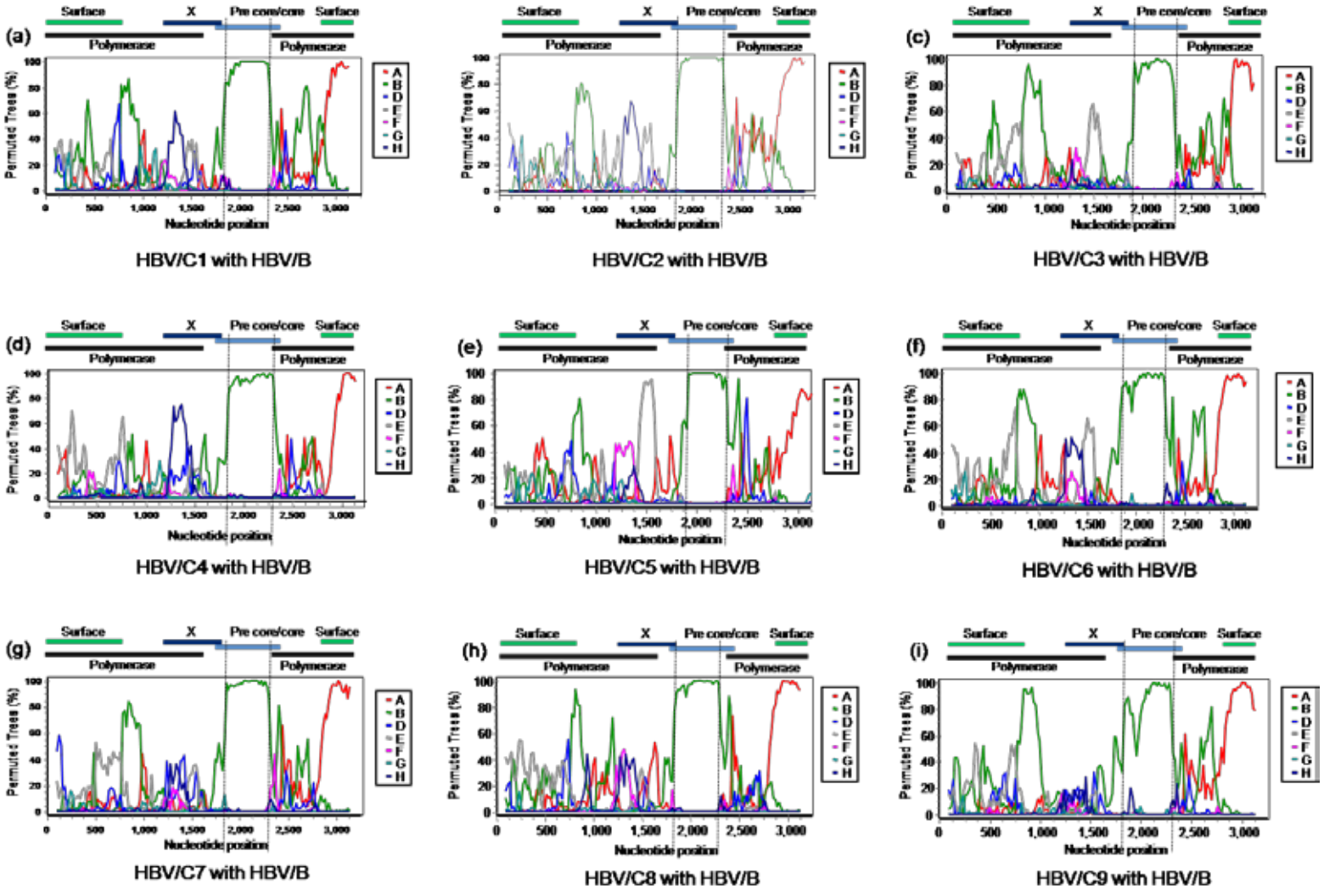
Amino acid sequence alignment of HBsAg of 179 HBV/C isolates. One hundred and seventy-two isolates of various HBV/C subgenotypes were examined for immune epitopes within the surface protein (HBsAg) including the subtype-determining amino acids 122 and 160. The isolates were retrieved from GenBank following their origins from various geographical regions of the East and Southeast Asia and the Papua-Pacific. Dots indicate amino acids identical to those of AF473543 of China used as consensus. The types and location of immune epitopes in HBsAg are indicated. Three unique substitutions (G18V, G44E, and L213I) separate all isolates into two clusters, the Asia and the Papua-Pacific. The substitution G18V is located in the class II HLA-DR-restricted T-cell epitope (residues 16-31) (Min *et al,* 1996), and the substitution G44E is located in the class I HLA-A2-restricted T-cell epitope (residues 41-49), coincident with the hotspot mutational domain (residues 28-51) of HBsAg (Tai *et al.,* 1997).

Supplementary Figure 3

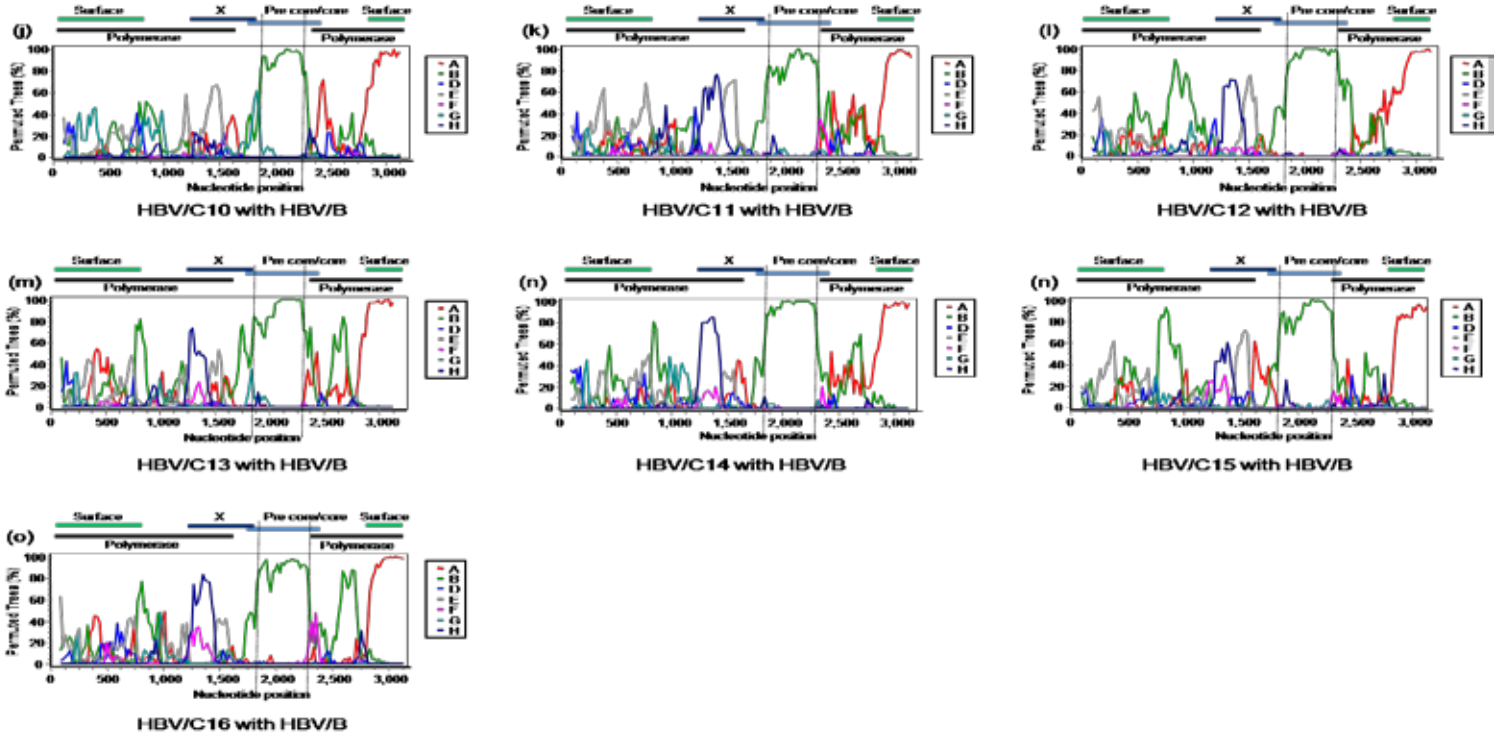


The amino acid sequence alignment of HBsAg 87 HBV/C isolates generated in this study. A total of 87 HBV/C isolates were collected from ethnically-defined hosts from various geographical regions of the Indonesian archipelago. Analysis was performed for variations of the amino acids of the surface protein (HBsAg) from residues 104-178. Dots indicate amino acids identical to those of JQ740646 of Minang (Sumatra) of Indonesia used as consensus. Shading indicates subtype-determining amino acids 122 and 160.

Supplementary Figure 4



Supplementary Figure 4. continued



Bootscan analysis of recombination breakpoints was examined for each HBV/C subgenotype groups (C1 to C16) with other 7 genotypes (HBV/A to HBV/H), and HBV/C1 to C16 used as the query, (a) to (o), respectively. Each HBV genotypes is analyzed over the complete genome using a 200 bp window size, 20 bp step size, 1.000 replicates, and neighbor joining method; each of these genotypes is represented by a different color as indicated. Vertical doted lines show the recombination breakpoints, and schematic diagrams of HBV genome is indicated above each figure. Recombination was detected in all HBV/C subgenotypes with genotype B in precore/core region from nt1820 to 2350 with various length of recombination.

CHAPTER 4

Occult hepatitis B in blood donors in Indonesia: altered antigenicity of the hepatitis B virus surface protein

Meta D. Thedja[1], Martono Roni[1], Alida R. Harahap[1,2], Nurjati C. Siregar[1,2], Susan I. Ie[1], David H. Muljono[1,§]

[1]Eijkman Institute for Molecular Biology, Jl. Diponegoro 69, Jakarta, Indonesia

[2]Faculty of Medicine, University of Indonesia, Jl. Salemba 6, Jakarta, Indonesia

[§]Corresponding author

Hepatol Int (2010) 4:608–614

Abstract

Background and aims: Occult hepatitis B virus infection (OBI) poses a challenge to the safety of blood donation. The prevalence of OBI is not well documented in Indonesia, while this information in such an endemic country is needed. This study was aimed to evaluate the prevalence of occult hepatitis B in blood donors from two cities of Indonesia, and to study the genetic variation and its effect on the predicted antigenicity of HBsAg.

Methods: Serum samples of 309 regular blood donors negative for HBsAg were tested for anti-HBs and anti-HBc. HBV DNA isolated from anti-HBc positive samples were analyzed by PCR, cloning and sequencing. Antigenic properties of identified HBsAg mutants were predicted by calculation of the antigenic index.

Results**:** Of the 309 HBsAg negative samples, anti-HBc was positive in 134 (43.4%) and HBV DNA was detected in 25 (8.1%). Seven of the viremic samples had nucleotide substitutions (A521G, A551T, C582T, and A562G) in the S gene, causing amino acid mutations (T123A, M133L, and T143M) in the '*a*' determinant of HBsAg that resulted in changes in the predicted antigenicity.

Conclusions**:** OBI was detected in blood donors' samples in Indonesia. Anti-HBc was shown to be a better screening parameter than HBsAg, however, it might consequence in loss of donors particularly in endemic countries. HBsAg detection failure in this study might be due to mutations altering the protein antigenicity and/or the low-level carriage of HBV.

Background

Chronic Hepatitis B Virus (HBV) infection continues to be a global public health problem that affects an estimated 360 million individuals [1]. Two thirds of these HBV carriers live in Asia-Pacific region where hepatitis B is the leading cause of chronic hepatitis, cirrhosis and hepatocellular carcinoma (HCC) [2]. It is of particular concern in Indonesia that belongs to a region with an intermediate-to-high level of hepatitis B endemicity [1].

One important mode of HBV transmission is through contaminated blood transfusion. The safety of blood donation has become an important issue since occult HBV infection (OBI) has been detected and could be transmitted to the recipients [3-4]. OBI is defined as the presence of HBV DNA in serum and/or liver without detectable HBsAg [5]. It is found in several conditions [5-6]: (a) recovery from past infection defined by the

presence of anti-HBs, (b) chronic hepatitis with surface gene escape mutants that are not or poorly recognized by current assays, (c) chronic carriage without any marker of HBV infection other than HBV DNA, referred to as seronegative, and (d) chronic carriage with HBsAg too low to be detected and recognized solely by the presence of anti-HBc.

The occurrence of OBI to a large extent depends on the prevalence of HBV infection in the general population. It is most common in regions where HBV infection is endemic [7-8]. Since the first evidence of OBI was reported in 1979, there has been continuous increase in the number of publications on OBI covering various areas of bio-medical and public health aspects [9]. Most of the publications came from countries with low endemicity. The prevalence of OBI is not well documented in Indonesia although such information is urgently needed. To explore the extent of this problem, this initial study was performed with aims to evaluate the prevalence of occult hepatitis B in blood donors from two cities in Indonesia, to analyze the genetic characteristics of HBV, and to study the effect of the genetic alteration of HBV DNA on the predicted antigenicity of HBsAg.

Materials and Methods

Study samples

A total of 309 serum samples of regular blood donors negative for HBsAg, anti-HCV and anti-HIV (aged 17-56, mean 28.97 ± 8.81 years; male/female: 273/36) were used for this study. The samples were obtained from a serological surveillance for the main transfusion-transmitted infections including hepatitis B, conducted by the Indonesian Red Cross Blood Transfusion Unit in two cities of Indonesia, Solo in Java and Medan in Sumatra islands. Informed consent for participation in this study was obtained from each blood donor. All samples were collected in year 2004-2005 and stored at -70°C until used. The study protocol was in accordance with and approved by the Eijkman Institute Research Ethics Commission (EIREC No. 24/2007).

Serological detection of HBV markers

Prior to this study, all regular blood donors from the two cities were tested by two immunoassay procedures, namely Murex HBsAg Version 3 (Abbott/Murex Biotech Ltd) for screening, and Auzyme® Monoclonal (Abbott Laboratories) for confirmation. The HBsAg negative samples were employed for this study and examined by enzyme immunoassays for anti-HBs (AUSAB EIA®, Abbott Laboratories) and total anti-HBc (HBV Core Antigen CORZYME®, Abbott Laboratories) according to the manufacturer's

instruction. Quantification of anti-HBs was performed using Ausab Anti-HBs Quantitation Panel (Abbott Laboratories). Anti-HBs concentration equal to or greater than 10 IU/L was considered positive.

HBV extraction and polymerase chain reaction (PCR)

HBV DNA was isolated from 100 μL of HBsAg-negative and anti-HBc-positive sera by proteinase-K digestion and phenol-chloroform-isoamyl alcohol extraction [10]. The resulting precipitate was resuspended in 20 μL of double-distilled water and stored at -20^{o} C. Nested PCR was performed targeting a segment within the S gene that codes for the *'a'* determinant of HBsAg. The outer primers were S2-1 and S1-2, while the inner primers were S088 and S2-2 [11-12]. Sequences of the oligonucleotide primers are shown in Table 1. Denaturizing, annealing and extension were done at 94°C for 30 s, 55°C for 30 s, and 72°C for 1 min, respectively, for both rounds (35 cycles for the first and 25 for the second steps of PCR).

Table 1. HBV-DNA specific primers used in nested PCR, clone analysis and sequencing

Primer	Nucleotide sequence (5' → 3')	Nucleotide position#	Polarity	Ref
S2-1	CAAGGTATGTTGCCCGTTTG	455 – 474	sense	28
S088	TGTTGCCCGTTTGTCCTCTA	462 – 471	sense	29
S1-2	GCCATTTGTTCAGTGGTTCG	685 – 704	antisense	28
S2-2	TGGCTCAGTTTACTAGTGCC	668 – 687	antisense	28
PS8-1	GTCACCATATTCTTGGGAAC	2817 – 2836	sense	30
HS6-2	GCCAAGTGTTTGCTGACGCA	1175 – 1194	antisense	
HS4-2	CCTATTGATTGGAAGGTGTG	970 – 989	antisense	30
T728	GGAATCAAACCTTATTATCC	2688 – 2707	sense	30
T703	CAGAGTCTAGACTCGTGGTG	242 – 261	antisense	30
PS5-2	CTCGTGTTACAGGCGGGGTT	191-210	antisense	
M13F	GCCAGGGTTTTCCCAGTCACGAC	2949 – 2972	sense	
M13R	GTCATAGCTGTTTCCTGTGTGA	176 - 197	antisense	

 # based on *EcoRI* site numbering

To confirm the detection of HBV DNA and mutations in the *'a'* determinant of the S gene, semi-nested PCR was also performed to amplify part of the overlapping P gene using primers PS8-1 and HS6-2 for the first round, and PS8-1 and HS4-2 for the second round (Table 1) [13]. To determine the presence of mutations in the S promoter that could

affect the transcription of the S gene, a fragment covering the PreS1 region of the S gene was amplified by semi nested PCR using primer set T728/T703 for the first round, and T728/PS5-2 for the second round, with primer sequences as shown in Table 1 [13].

The amplification products were visualized on ethidium bromide-stained 2 percent agarose gel under ultraviolet light. Kwok and Higuchi rules were followed strictly in all experimental steps [14]. The nucleotide positions of the primers used in this study are based on *EcoRI* site. Positive PCR products were purified using QIAquick™ PCR Purification kit (QIAGEN, Hilden, Germany).

Cloning and sequence analysis of S gene

Purified PCR products of S gene were ligated to pGEM®T Easy vector (Promega Co., Madison, WI) and transformed into *E.coli* JM109. Transformed bacteria were selected by plating on Luria-Bertoni agar in the presence of ampicillin at 100 µg/mL and screened by PCR using primers S088 and S2-2. Six clones of each sample containing the HBV insert were selected and grown overnight in Luria-Bertoni broth containing 50 µg/mL ampicillin. Recombinant plasmids were recovered by standard alkaline lysis miniprep procedure and sequenced using the Bigdye Terminator Cycle Sequencing Ready Reaction kit (Applied Biosystems, Foster City, CA) on an automatic sequencer (Applied Biosystems 337 DNA, Perkin Elmer). Primers M13F and M13R were used as the sequencing primers (Table 1). Each clone was sequenced bidirectionally by two independent reactions. The nucleotide sequences comprising of 226 bp fragment of the S gene were aligned and compared using software BioEdit Sequence Alignment Editor Ver. 7.0.5.2 [15] with two wild type HBV sequences, M54923 (Genotype B, *adw*) and AP011097 (Genotype C, *adr*) retrieved from GenBank [16-17].

Direct sequencing of P gene and Pre-S1 region of S gene

Direct sequencing of purified PCR products of the P gene and Pre-S1 region of the S gene was also performed using the same methods. Each sample was also sequenced bidirectionally using PCR products from two independent reactions. The sequences obtained were aligned and compared with that of M54923 retrieved from Genbank.

Calculation of antigenicity and secondary structure

To determine whether changes in amino acid sequence alter the antigenicity of HBsAg, a study of Jameson-Wolf Antigenic Index Prediction was performed using Lasergene Protean v8.1 program (DNASTAR Inc., Madison, WI). The antigenicity index prediction

study combines information from primary amino acid sequence hydrophilicity (Hoop-Wood and Kyte Doolittle method), surface probability (Emini method), and backbone flexibility (Karplus-Schultz method) predictions, together with secondary structure predictions (Chou-Fasman and Robson-Garner) [18-19].

Results

Detection of occult HBV in blood donors

Of 309 HBsAg negative blood donor samples, 134 (43.4%) were positive for total anti-HBc, referred to hereafter as anti-HBc, and therefore had serological evidence of prior and/or ongoing HBV infection. The remaining samples were negative for anti-HBc and excluded from this study. Of the 134 samples with anti-HBc, 68 (50.7%) had detectable anti-HBs and 66 (49.3%) were anti-HBs negative (referred as to isolated anti-HBc). HBV DNA was detected in 25 (18.7%) of all anti-HBc positive samples, including 6 (8.8%) from the anti-HBs- positive/anti-HBc-positive and 19 (28.8%) from the isolated anti-HBc groups. Thus, the overall prevalence of OBI in 309 blood donors was 8.1%. Of these, 6 (24%) were anti-HBs positive with 2 samples had antibody titer greater than 100 IU/L. Demographic and serologic data of donors with HBV DNA are shown in Table 2.

Sensitivity of the nested PCR performed in this study was validated using a panel of sera with various HBV DNA titers tested by COBAS TaqMan 48 Real-Time PCR (Roche Molecular System, Branchburg, NJ, USA). The nested PCR method was capable of detecting HBV DNA at titers lower than the detection limit of the COBAS TaqMan 48 Real-Time PCR (6 IU/mL), thus, met the sensitivity requirement for detection of occult hepatitis B of less than 10 IU/mL [5].

Analysis of nucleotide sequences and protein products

Of the 150 clones derived from 25 HBV DNA positive samples, 42 (28%) from 7 samples had nucleotide substitutions: 6 clones from 1 sample exhibiting A521G substitution, 6 clones from 1 sample with A551T and A562G substitutions, and the other 30 clones from 5 samples had C582T substitution.

Table 2. Serological markers of HBV DNA positive blood donors and mutation pattern of HBsAg

No.	Donor ID	Age (yrs)	Sex (M/F)	HBsAg	Anti-HBc	Anti-HBs	Anti-HBs titer (IU/L)	HBV-DNA	HBsAg mutation
1	**2013**	45	L	-	+	+	346	+	T143M
2	**2509**	41	L	-	+	+	103	+	wt
3	**2314**	20	L	-	+	+	59	+	wt
4	**2096**	45	P	-	+	+	46	+	wt
5	**2050**	22	L	-	+	+	39	+	wt
6	**2542**	38	L	-	+	-	10	+	wt
7	**2350**	25	L	-	+	-	9	+	M133L
8	**2411**	22	L	-	+	-	9	+	wt
9	**2028**	23	L	-	+	-	8	+	wt
10	**2054**	25	L	-	+	-	8	+	wt
11	**2072**	20	L	-	+	-	8	+	T143M
12	**2362**	35	L	-	+	-	8	+	wt
13	**2407**	27	L	-	+	-	8	+	wt
14	**2412**	23	L	-	+	-	8	+	wt
15	**2537**	29	L	-	+	-	8	+	wt
16	**2083**	28	L	-	+	-	7	+	T143M
17	**2361**	37	L	-	+	-	7	+	T123A
18	**2414**	21	L	-	+	-	7	+	wt
19	**2427**	30	L	-	+	-	7	+	wt
20	**2182**	32	L	-	+	-	6	+	T143M
21	**2392**	35	L	-	+	-	5	+	wt
22	**2524**	26	L	-	+	-	5	+	wt
23	**2357**	25	L	-	+	-	4	+	wt
24	**2133**	28	L	-	+	-	3	+	T143M
25	**2351**	23	L	-	+	-	3	+	wt

*** A level of anti-HBs equal to or higher than 10 IU/L was considered positive.**

Three of these four substitution patterns – A521G, A551T, and C582T – caused mutations within the *'a'* determinant: T123A in 1 sample, M133L in 1 sample, and T143M in 5 samples, respectively, whereas pattern A562G caused silent mutation. The HBsAg mutation patterns identified in these HBV DNA positive samples are shown in Table 2 and Figure 1. As a consequence of HBV overlapping open reading frames, the nucleotide changes in the S gene (A521G, A551T, and A562G) were associated with amino acid alterations in the reverse-transcriptase domain of HBV polymerase protein: rtN131S, rtY141F, and rtM145V, respectively, while C582T caused silent mutation [20]. The remaining 108 (72%) clones from 18 samples had identical nucleotides to that of the M54923 sequence, even when the screening was extended to 10 additional clones from

each sample. The 42 nucleotide sequences from 7 samples with substitution and representations from each of the 18 samples with wild type HBV were deposited at GenBank (AN EF507434-EF507475 and HM116516- HM116533). Analysis of preS1 sequences generated from direct sequencing showed wild type HBV in all samples.

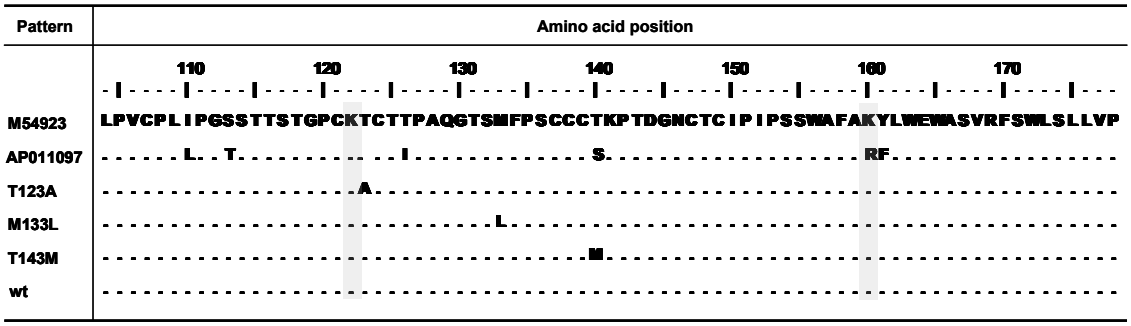


Figure 1. Mutation pattern of the hepatitis B surface antigen (HBsAg) of 25 blood donor samples with HBV DNA. M54923 (Genotype B, *adw*) and AP011097C (C, *adr*) obtained from GenBank are used as reference sequences. Dots (·) represent amino acids identical to those in the M54923. Amino acids that determine HBsAg serotype are shaded. Three substitution patterns were observed: T123A (in 1 donor), M133L (in 1 donor), and T143M (in 5 donors). *wt* represents sequences with no amino acid substitution in 18 donors.

Alteration in predicted antigenicity of HBsAg

To determine the effect of amino acid changes to HBV surface protein antigenicity, a Jameson & Wolf antigenic index prediction was conducted [19]. All mutation patterns found in this study showed changes of the secondary structure and the predicted antigenicity of HBsAg (Figure 2). Sequence with T123A pattern, a substitution located one residue upstream of the first HBsAg loop, only experienced relatively minor and localized change in the antigenicity index at position 123 and its close proximity compared to the wild type. In M133L pattern, with substitution in the first loop of HBsAg in the secondary structure prediction, the effect was also slight and localized (-0.05 instead of -0.2 at residue 133). However, sequences with T143M pattern, the most prevalent variant found in this study, showed significant decrease in the pattern of antigenicity of the second HBsAg loop, which was observed in a larger area surrounding the amino acid substitution (Figure 2).

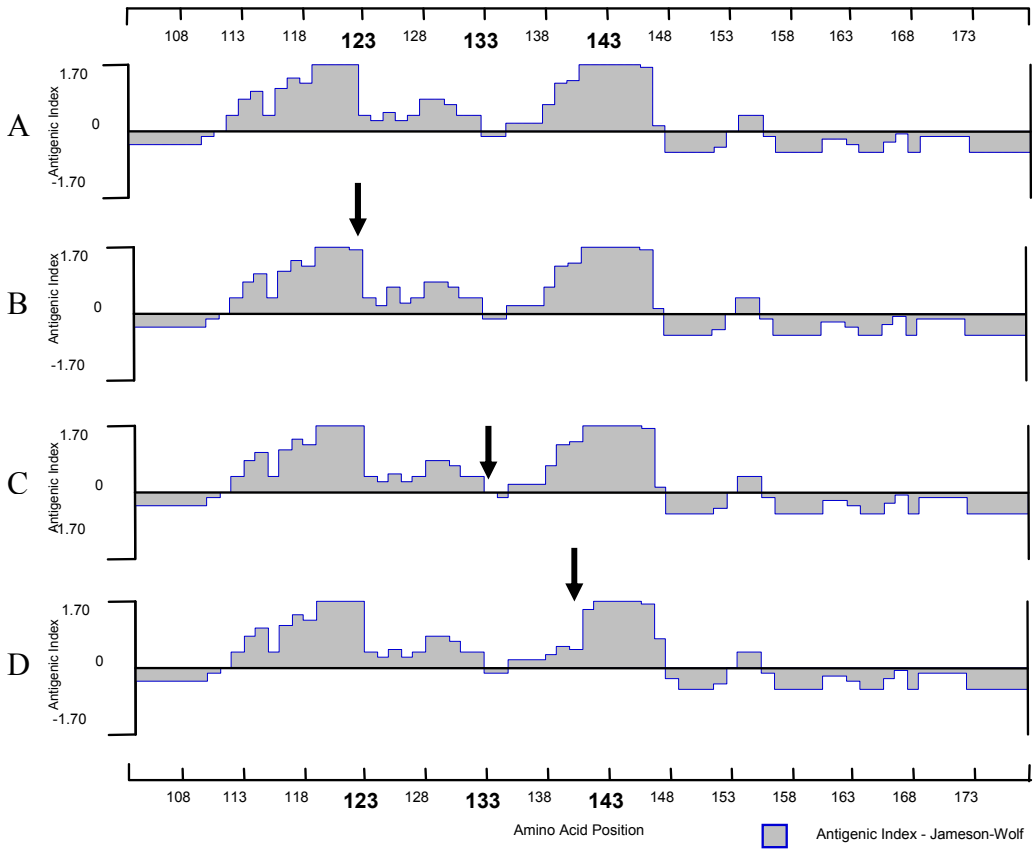


Figure 2. Antigenicity plots based on amino acid sequences of the *'a'* determinant region of HBsAg. Antigenicity plots of the reference sequence (M54923) (A) and of HBV mutants isolated from blood donors found in this study, with amino acid changes T123A (B), M133L (C), and T143M (D). The change in each antigenic index is indicated by arrow, with the most significant alteration observed in the T143M substitution while T123A and M133L show minor antigenicity changes.

Discussion

The implications of HBV infection involve several clinical aspects, varying from chronic asymptomatic carriers to complicated liver diseases including liver cirrhosis and hepatocellular carcinoma (HCC) [1-2]. OBI has appeared to have similar infectivity and pathogenicity in the development of fulminant hepatitis, liver cirrhosis and HCC [7], and would possibly affect the safety of blood transfusion [21].

OBI is also related to the endemicity of HBV infection. It is most commonly reported in high endemic areas and infrequently detected in low endemic areas [7-8]. In this study, OBI was detected in 8.1% of regular blood donors' samples, nearly equals to the average 9.4% HBsAg prevalence in Indonesia [22]. It was higher than those found elsewhere (7% in Taiwan and 1.4% in Ghana) [21,23]. This evidence alarms us that OBI in blood donors with negative HBsAg status is not negligible.

The detection of anti-HBc in HBsAg-negative individuals has been considered a marker of past HBV exposure and/or of resolved infection. However, application of molecular biology techniques has shown that HBV viremia are detectable in 1.33 to 38 % of HBsAg-negative/anti-HBc-positive donors [24-25]. In this study, HBV DNA was detected in 25 (18.7%) of anti-HBc-positive regular blood donors with higher frequency of HBV DNA in isolated anti-HBc subjects than in those with anti-HBc and anti-HBs. This finding highlights the importance of anti-HBc compared to other serological HBV markers for predicting latent HBV infection in apparently healthy individual, and reiterates that the implementation of anti-HBc screening would improve the safety of blood supply [25-27]. However, in highly-endemic regions including Indonesia, anti-HBc screening would be impractical due to the high loss of potential donors (approx. 70% of isolated anti-HBc donors). This study could give support to the potential use of molecular detection as an alternative when it has become widely available at lower cost for public health.

The presence of anti-HBs and anti-HBc is usually indicative of immunity after infection. In some countries such as Germany, Austria and Japan, blood units with anti-HBs levels greater than 100 IU/L is considered to be safe [26]. However, there was evidence that transmission of HBV from occult hepatitis B subjects occurred in the presence of concurrent neutralizing anti-HBs in the same specimen [28]. Detection of HBV DNA in some anti-HBs positive samples in this study raises doubt whether the absence of HBsAg together with the presence of anti-HBs could reflect the safety of blood donations. A similar report from Italy also supports the notion that some blood donors with anti-HBs titer over 100 IU/L still had detectable HBV DNA [8]. Overall, these results raise several important public health issues: the absence of HBsAg as an HBV infection

marker does not exclude the possibility of viral transmission; anti-HBc positive-sera potentially contain HBV; and the presence of anti-HBs is not a sign of total HBV eradication.

In this study, although the frequently emerged variant G145R was not found, some other variants showing single amino acid substitution patterns within the *'a'* determinant that had been reported previously were observed: T123A, M133L, and T143M [29-31]. All isolates showing T123A and M133L substitutions, together with 4 of T143M isolates, were obtained from isolated anti-HBc samples; while one of the T143M isolates was found in samples with anti-HBc and anti-HBs. This finding might suggest that mutation in the *'a'* determinant is more frequently observed in the isolated anti-HBc samples (Table 2).

Conformationally-dependent antigenic determinant might be affected by changes of its amino acid residues [32-33]. While M133L substitution did not show significant alteration, both T123A and T143M mutations demonstrated results that should be taken into account. Pattern T123A resided in close proximity to the cysteine residues at 121 and 124, which form disulfide bonds that are important for maintaining the *'a'* determinant's conformation [30,33]. This close proximity substitution might cause alteration in the steric hindrance that disturbs the disulfide bonds, and hence affect HBsAg conformation and its detection. The other substitution, T143M, caused marked alterations demonstrated by extensive changes of antigenic index of the mutated amino acid and its surroundings. This pattern affected the second loop of the *'a'* determinant, which is more antigenic than the first [32,34]. Thus, mutations in the second loop would be more significantly disrupting the HBsAg antigenicity. These substitutions might partly explain the detection failure of HBsAg in this study.

It is acknowledged that this conclusion is based on mathematical modelling and may not reflect actual changes in antibody recognition. Further protein model prediction based on these amino acid substitution patterns might explain the conformational changes of HBsAg, and assays to confirm reduction in binding affinity of the altered epitopes to monoclonal anti-HBs are suggested.

As a consequence of gene overlapping, the nucleotide substitutions also caused amino acid mutations in the HBV polymerase: rtN131S, rtY141F, and rtM145V. These mutations lay between domain A and B of the reverse transcriptase region of polymerase protein, which is crucial for its function in HBV replication processes [35-36]. One or several of the identified mutations could be responsible for diminished rate of replication causing the detection failure of HBsAg.

In this study, 72% of the viremic donors had HBV DNA with wild type S gene. This finding could indicate that there are other factors beside antigenic property that may cause negative HBsAg status. Since all samples were negative for anti-HCV and had HBV of the same genotype and serotype, low viral titer would provide an alternative explanation for the wild type HBV undetectable by HBsAg serological assay. This is supported by the fact that HBV DNA in these samples was detectable only in the second-round of nested PCR.

In conclusion, occult Hepatitis B was detected in samples of regular blood donors from Indonesia. Failure of HBV detection in these cases demonstrated that anti-HBc seemed to be a better screening parameter than HBsAg. Amino acid mutation in the S gene which altered HBsAg antigenic property might in part be the molecular background of the failure of HBsAg detection. Other factor contributing to the insensitivity of the assay could be the low titer of viral load. Further experimental studies are needed to confirm the changes in antigenicity of these HBsAg variants. Studies involving more samples from various regions in Indonesia are important in investigating the magnitude of occult hepatitis B infection and the characteristics of occult HBV strains among blood donors in Indonesia.

References

1. World Health Organization: Hepatitis B Vaccines: Weekly epidemiological record. WHO annual report. Switzerland; 2009, 84:405-420.
2. Liaw YF, Leung N, Kao JH, Piratvisuth T, Gane E, Han KH, Guan R, Lau GKK, Locarnini S. Asian-Pacific consensus statement on the management of chronic hepatitis B: a 2008 update. Hepatol Int 2008;2:263–283.
3. Hoofnagle JH, Seeff LB, Bales ZB, Zimmerman HJ. Type B hepatitis after transfusion with blood containing antibody to hepatitis B core antigen. N Engl J Med 1978;298:1379-1383.
4. Levicnik-Stezinar S, Rahne-Potokar U, Candotti D, Lelie N, Allain JP. Anti-HBs positive occult hepatitis B virus carrier blood infectious in two transfusion recipients. J Hepatol 2008;48:1022-1025.
5. Raimondo G, Allain JP, Brunetto MR, Buendia MA, Chen DS, Colombo M, Craxi A, Donato F, Ferrari C, Gaeta GB, Gerlich WH, Levrero M, Locarnini S, Michalak T, Mondelli MU, Pawlotsky JM, Pollicino T, Prati D, Puoti M, Samuel D, Shouval D, Smedile A, Squadrito G, Trepo C, Villa E, Will H, Zanetti AR, Zoulim F. Statements from the *Taormina* expert meeting on occult hepatitis B virus infection. J Hepatol 2008;49:652-657.
6. Allain JP. Occult hepatitis B virus infection: implications in transfusion. Vox Sang 2004;86:83-91.

7. Torbenson M, Thomas DL. Occult hepatitis B. Lancet Infect Dis 2002;2:479-486.

8. Manzini P, Girotto M, Borsotti R, Giachino O, Guaschino R, Lanteri M, Testa D, Ghiazza P, Vacchini M, Danielle F, Pizzi A, Valpreda C, Castagno F, Curti F, Magistroni P, Abate ML, Smedile A, Rizzetto M. Italian blood donors with anti-HBc and occult hepatitis B infection. Haematologica 2007;92:1664-1670.

9. Chemin I, Trepo C. Clinical impact of occult HBV infections. J Clin Virol 2005;34(Suppl 1):S15-S21.

10. Okamoto H, Yano K, Nozaki Y, Matsui A, Miyazaki H, Yamamoyo K, Tsuda F, Machida A, Mishiro S. Mutation within the S gene of Hepatitis B Virus transmitted from mothers to babies immunized with Hepatitis B Immunoglobulin and vaccine. Pediatric Res 1992;32:264-268.

11. Iizuka H, Ohmura K, Ishijima A, Satoh K, Tanaka T, Tsuda F, Okamoto H, Miyakawa Y, Mayumi M. Correlation between anti-HBc titers and HBV DNA in blood units without detectable HBsAg. Vox Sang 1992;63:107–111.

12. Okamoto H, Nishizawa T. Non-B Non-C Non-G Hepatitis Virus Gene, Polynucleotide, Polypeptide, Virion, Method for Separating Virion, and Method for Detecting Virus. 1992. http://www.freepatentsonline.com/EP1010759.html (with permission).

13. Takahashi K, Akahane Y, Hino K, Ohta Y, Mishiro S. Hepatitis B virus genomic sequence in the circulation of hepatocellular carcinoma patients: comparative analysis of 40 full-length isolates. Arch Virol 1998;143:2313-2326.

14. Kwok S, Higuchi R. Avoiding false positives with PCR. Nature 1989;339:237-238.

15. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp 1999;41:95-98.

16. Sastrosoewignjo RI, Omi S, Okamoto H, Mayumi M, Rustam M, Sujudi T. The complete nucleotide sequence of HBV DNA clone of subtype adw (pMND122) from Menado in Sulawesi Island, Indonesia. ICMR Ann 1987;7:51-60.

17. Mulyanto, Depamede SN, Surayah K, Tsuda F, Ichiyama K, Takahashi M, Okamoto H. A nationwide molecular epidemiological study on hepatitis B virus in Indonesia: identification of two novel subgenotypes, B8 and C7. Arch Virol 2009;154(7):1047-1059.

18. Jameson BA, Wolf H. The antigenic index: a novel algorithm for predicting antigenic determinants. Cabios 1988;4:181-186.

19. Ghany MG, Ayola B, Villamil FG, Gish RG, Rojter S, Vierling JM, Lok AS. Hepatitis B Virus mutants in liver transplant recipients who were reinfected despite Hepatitis B Immune Globulin prophylaxis. Hepatology 1998;27:213-221.

20. Stuvyer LJ, Locarnini SA, Lok A, Richman DD, Carman WF, Dienstag JL, Schinazi RF, The HEP DART IC. Nomenclature for antiviral-resistant human Hepatits B Virus mutations in the Polymerase region. Hepatology 2001;33:751-757.

21. Wang JT, Lee CZ, Chen PJ, Wang TH, Chen DS. Transfusion-transmitted HBV infection in an endemic area: the necessity of more sensitive screening for HBV carriers. Transfusion 2002;42:1592-1597.

22. Khan M, Dong JJ, Acharya SK, Dhagwahdorj Y, Abbas Z, Jafri SMW, Mulyono DH, Tozun N, Sarin SK. Hepatology issues in Asia: Perspectives from regional leaders. J Gastroenterol Hepatol 2004;19:S419-S430.

23. Owusu-Ofori S, Temple J, Sarkodie F, Anokwa M, Candotti D, Allain JP. Predonation screening of blood donors with rapid test: implementation and efficacy of a novel approach to blood safety in resource-poor settings. Transfusion 2005;45:133-140.

24. Yotsuyanagi H, Yasuda K, Moriya K, Shintani Y, Fujie H, Tsutsumi T, Nojiri N, Juji T, Hoshino H, Shimoda K, Hino K, Kimura S, Iino S, Koike K. Frequent presence of HBV in the sera of HBsAg-negative, anti-HBc-positive blood donors. Transfusion 2001;41(9):1093-1099.

25. El-Sherif AM, Abou-Shady MA, Al-Hiatmy MA, Al-Bahrawy AM, Motawea EA. Screening for hepatitis B virus infection in Egyptian blood donors negative for hepatitis B surface antigen. Hepatol Int 2007;1:469-470.

26. Hollinger FB. Hepatitis B virus infection and transfusion medicine: science and the occult. Transfusion 2008;48:1001-1026.

27. Niederhauser C, Taleghani BM, Graziani M, Stolz M, Tinguely C, Schneider P. Blood donor screening: how to decrease the risk of transfusion-transmitted hepatitis B virus? Swiss Med Wkly 2008;138(9-10):134-141.

28. Matsumoto C, Tadokoro K, Fujimura K, Hirakawa S, Mitsunaga S, Juji T. Analysis of HBV infection after blood transfusion in Japan through investigation of a comprehensive donor specimen repository. Transfusion 2001;41:878-884.

29. Ho MS , Mau YC, Lu CF, Huang SF, Hsu LC, Lin SR, Hsu HM. Patterns of circulating hepatitis B surface antigen variants among vaccinated children born to hepatitis B surface antigen carrier and non-carrier mothers: A population-based comparative study. J Biomed Sci 1998;5:355-362.

30. Hou JL, Wang ZH, Cheng JJ, Lin YL, Lau GKK, Sun J, Zhou FY, Waters J, Karayiannis P, Luo KX. Prevalence of Naturally Occuring Surface Gene Variants of Hepatitis B Virus in Non-immunized Surface Antigen-Negative Chinese Carriers. Hepatology 2001;34:1027-1034.

31. Thakur V, Kazim SN, Guptan RC, Hasnain SE, Bartholomeusz A, Malhotra V, Sarin SK. Transmission of G145R Mutant of HBV to an Unrelated Contact. J Med Virol 2005;76:40-46.

32. Carman WF, Zanetti AR, Karayiannis P, Waters J, Manzillo G, Tanzi E, Zuckerman AJ, Thomas HC. Vaccine-induced escape mutant of hepatitis B virus. Lancet 1990;336:325-329.

33. Seddigh-Tonekaboni S, Waters JA, Jeffers S, Gehrke R, Ofenloch B, Horsch A, Hess G, Thomas HC, Karayiannis P. Effect of variation in the common “a” determinant on the antigenicity of hepatitis B surface antigen. J Med Virol 2000;60:113–121.

34. Zuckerman AJ. Effect of hepatitis B virus mutants on efficacy of vaccination. Lancet 2000;355:1382-1383.

35. Radziwill G, Tucker W, Schaller H. Mutational Analysis of the Hepatitis B Virus P Gene Product: Domain Structure and RNase H Activity. J Virol 1990;64(2):613-620.

36. Torresi J, Earnest-Silveira L, Civitico G, Walters TE, Lewin SR, Fyfe J, Locarnini SA, Manns M, Trautwein C, Bock TC. Restoration of Replication Phenotype of Lamivudine-Resistant Hepatitis B Virus Mutants by Compensatory Changes in the "Fingers" Subdomain of the Viral Polymerase Selected as a Consequence of Mutations in the Overlapping S Gene. Virology 2002;299:88-99.

Erratum to: Occult hepatitis B in blood donors in Indonesia: altered antigenicity of the hepatitis B virus surface protein

Meta D. Thedja[1], Martono Roni[1], Alida R. Harahap[1,2], Nurjati C. Siregar[1,2], Susan I. Ie[1], David H. Muljono[1,§]

Hepatol Int (2010) 4:788

In the above–mentioned article,

1. There are some missing sentences in the section “Background” of the paper. The last two sentences should read as below:

The prevalence of OBI is not well documented in Indonesia although such information is urgently needed. To explore the extent of this problem, this initial study was performed with aims to evaluate the prevalence of occult hepatitis B in blood donors from two cities in Indonesia, to analyze the genetic characteristics of HBV, and to study the effect of the genetic alteration of HBV DNA on the predicted antigenicity of HBsAg.

2. Table 1: In the column “References”, the numbers of the references were incorrectly stated. The correct Table 1 should have appeared as shown below

Table 1 HBV DNA-specific primers used in nested PCR, clone analysis, and sequencing

Primer	Nucleotide sequence (5’ → 3’)	Nucleotide position#	Polarity	Ref
S2-1	CAAGGTATGTTGCCCGTTTG	455 – 474	sense	11
S088	TGTTGCCCGTTTGTCCTCTA	462 – 471	sense	12
S1-2	GCCATTTGTTCAGTGGTTCG	685 – 704	antisense	11
S2-2	TGGCTCAGTTTACTAGTGCC	668 – 687	antisense	11
PS8-1	GTCACCATATTCTTGGGAAC	2817 – 2836	sense	13
HS6-2	GCCAAGTGTTTGCTGACGCA	1175 – 1194	antisense	
HS4-2	CCTATTGATTGGAAGGTGTG	970 – 989	antisense	13
T728	GGAATCAAACCTTATTATCC	2688 – 2707	sense	13
T703	CAGAGTCTAGACTCGTGGTG	242 – 261	antisense	13
PS5-2	CTCGTGTTACAGGCGGGGTT	191-210	antisense	
M13F	GCCAGGGTTTTCCCAGTCACGAC	2949 – 2972	sense	
M13R	GTCATAGCTGTTTCCTGTGTGA	176 - 197	antisense	

 # based on *EcoRI* site numbering

CHAPTER 5

Prediction of conformational changes by single mutation in the hepatitis B virus surface antigen (HBsAg) identified in HBsAg-negative blood donors

Susan I. Ie[1*], Meta D. Thedja[1*], Martono Roni[1], David H. Muljono[1§]

[1]Eijkman Institute for Molecular Biology, Jl. Diponegoro 69, Jakarta, Indonesia

*These authors contributed equally to this work

§Corresponding author

Virology Journal 2010, 7:326

Abstract

Background

Selection of hepatitis B virus (HBV) by host immunity has been suggested to give rise to variants with amino acid substitutions at or around the *'a'* determinant of the surface antigen (HBsAg), the main target of antibody neutralization and diagnostic assays. However, there have never been successful attempts to provide evidence for this hypothesis, partly because the 3D structure of HBsAg molecules has not been determined. Tertiary structure prediction of HBsAg solely from its primary amino acid sequence may reveal the molecular energetic of the mutated proteins. We carried out this preliminary study to analyze the predicted HBsAg conformation changes of HBV variants isolated from Indonesian blood donors undetectable by HBsAg assays and its significance, compared to other previously-reported variants that were associated with diagnostic failure.

Results

Three HBV variants —T123A, M133L and T143M— and a wild type sequence were analyzed together with frequently emerged T123N, M133I, M133T, M133V, and T143L variants. Based on the Jameson-Wolf algorithm for calculating antigenic index, the first two amino acid substitutions resulted in slight changes in the antigenicity of the *'a'* determinant, while all four of the comparative variants showed relatively more significant changes. In the T143M pattern, changes in antigenic index were more significant, both in its coverage and magnitude, even when compared to T143L variant. These data were also partially supported by the tertiary structure prediction, in which the T143M pattern showed larger shift in the HBsAg second loop structure compared to the others.

Conclusions

Single amino acid substitutions within or near the *'a'* determinant of HBsAg may alter antigenicity properties of variant HBsAg, which can be shown by both its antigenic index and predicted 3D conformation. Findings in this study emphasize the significance of T143M variant, the prevalent isolate with highest degree of antigenicity changes found in Indonesian blood donors. This highlights the importance of evaluating the effects of protein structure alterations on the sensitivity of screening methods being used in detection of ongoing HBV infection, as well as the use of vaccines and immunoglobulin therapy in contributing to the selection of HBV variants.

Background

Hepatitis B Virus (HBV), the etiology of hepatitis B, is a DNA virus that replicates via an RNA intermediate [1]. It has a small partially double-stranded DNA genome of approximately 3.2 kilobases that contains four overlapping open reading frames, including one that encodes for the hepatitis B surface antigen (HBsAg) protein [1]. Diagnosis and screening of HBV infection is most commonly done by detection of the HBsAg by means of antibody-based assays [2]. These assays target the *'a'* determinant, the highly homologous region within HBsAg, which is also used as the main target of antibody generated by hepatitis B vaccines [2]. However, there have been reports on the failure of these assays in detecting HBsAg in infected individuals, which include inactive HBV carriers, vaccinated children born to mothers with HBV infection, and liver transplant recipients treated with hepatitis B immunoglobulin (HBIg) therapy [3-5].

Recognition of the *'a'* determinant by antibody against HBsAg (anti-HBs) depends on its 3D conformation, which also relies on the amino acid sequence of the regions flanking the *'a'* determinant [6-7]. To date, there have never been successful attempts on crystallizing native HBsAg molecules for structure determination purposes. Tertiary structures of HBsAg have not been fully determined, aside from its nature as a membrane spanning protein with four trans-membrane helices and a major hydrophilic region that is exposed on the surface of the virus [7-8]. It is of interest to be able to predict the tertiary structure of HBsAg solely from its primary amino acid sequence, because pathogen recognition by the host immune system is mainly based on protein-protein interaction, which depends on the conformation of the interacting proteins. We carried out this preliminary study to analyze the prediction of HBsAg conformation changes as caused by variations in the S gene of HBV isolated from Indonesian HBsAg-negative blood donors in comparison with variants frequently reported from various regions of the world. The results of this study may contribute in better understanding the host-pathogen interaction as well as paving the way to develop better techniques in designing diagnostic tools and vaccine candidates for hepatitis B.

Materials and Methods

Sample selection and preparation

This study is part of a larger project investigating the main transfusion-transmitted infections including hepatitis B in regular blood donors by the Indonesian Red Cross in

two cities of Indonesia, Medan of Sumatra and Solo of Java islands. Previous study by Thedja *et al*., 2010 showed that HBV DNA was detected in 25 (8.1%) of 309 HBsAg-negative blood donors [9]. HBV DNA in the blood donors' samples was undetectable by quantitative PCR and detectable only in the second-round of nested PCR, which was capable of detecting HBV DNA at titres lower than the detection limit of the Cobas-Taqman 48 Real-Time PCR (Roche Molecular System, Branchburg, NJ, USA), 6 IU/mL [9-10]. The sequences of HBV DNA isolated in the study had been deposited in GenBank under Accession Nos**.** EF507434-EF507475 and HM116516-HM116533. To analyze the HBsAg conformation changes resulted from variations in the S gene, we first aligned the translated nucleotide sequences of HBV isolated from the Indonesian HBsAg-negative blood donors with a wild type reference (M54923; genotype B/*adw*) retrieved from GenBank [11], using BioEdit Sequence Alignment Editor Ver. 7.0.5.2 software [12]. Next, we searched for more HBV variants reported in association with medical and public health issues (problems in diagnostic assays and/or escape to vaccine/HBIg therapy) from published articles and GenBank database, focusing on variants with substitutions at the corresponding amino acid positions. Totally, an additional 5 sequences were retrieved and analyzed for their antigenic index calculation.

Prediction of antigenicity

Translated HBsAg sequences that contain mutations were analyzed with Jameson-Wolf algorithm in the Lasergene Protean v8.1 program (DNASTAR Inc., Madison, WI) to predict the antigenic index of each consensus sequence. This algorithm integrates several parameters to calculate the antigenicity of the sequence based on the characteristics of its primary amino acid chain: hydrophilicity (Hopp-Woods), surface probability (Emini), flexibility of the protein backbone (Karplus-Schulz), and secondary structure prediction (Chou-Fasman and Garnier) using the following equation [13]:

$$A_i = \sum_{i=1}^{N} 0.3\,(H_i) + 0.15\,(S_i) + 0.15\,(F_i) + 0.2\,(CF_i) + 0.2\,(RG_i)$$

$$A_i = \sum_{i=1}^{N} 0.3\,(H_i) + 0.15\,(S_i) + 0.15\,(F_i) + 0.2\,(CF_i) + 0.2\,(RG_i)$$

with regions of positive A_i value clusters indicate possible antigenic determinant.

Tertiary structure prediction

Based on structural alignment using Template Identification tool from Swiss-Model by InterPro Scan, BLASTP 2.2.9, PSI-BLAST, and HHSEARCH v. 1.5.01 software [14-17], no template structure was found in ExPDB template library for the 226-amino-acid-long

HBsAg [18]. Therefore, tertiary structures of the HBsAg variants found in Indonesian blood donor samples were predicted using free modelling, or often termed as '*ab initio*' or '*de novo*' modelling [19]. In this study, we used I-TASSER method, a protein structure modelling approach based on an algorithm consists of consecutive steps of threading, fragment assembly, and iteration to obtain structure with the lowest energy as described previously [20-22]. All structure predictions of wild type reference sequence and the variants were predicted separately using individual I-TASSER queries, and visualized using DeepView/Swiss-PdbViewer [23].

Results

Characterization of HBV mutants

Sequencing of partial HBV surface gene of the clones derived from 25 HBV DNA positive samples [9] showed nucleotide substitutions in 7 samples: A521G in one sample, A551T and A562G in one sample, and C582T in five samples. Of the four nucleotide substitutions, three single mutation patterns (T123A, M133L and T143M) of HBV surface protein were observed, while A562G was found to be a nonsense mutation. These mutation positions corresponded with those of five isolates known to be associated with problem in diagnostic assays and/or escape to vaccine/HBIg therapy: T123N, M133I, M133T, M133V, and T143L [5,24-29] (Fig.1). The remaining 18 (72%) samples did not show any nucleotide substitutions [9].

Prediction of antigenicity

Prediction of antigenic index of mutant sequences notably revealed altered antigenicity at and around the sites of amino acid substitutions compared to the wild type sequence (Table 1). In T123A substitution, several amino acids were affected by this single substitution. Antigenic index values of four amino acids at the region around amino acid position 123 was altered between -0.4 to +0.2 in magnitude. In contrast, only a small antigenicity change was detected (from -0.2 to -0.05) at the single amino acid site of M133L substitution. Most significant changes were observed in the T143M substitution. In this last pattern, antigenic index of the residues at position 143 and up to 5 amino acids both upstream and downstream of this site were observed to be altered between -1.07 to +0.62 in magnitude. These antigenic index changes were grouped into collectively negative alterations – i.e. more hydrophobic characteristics – upstream of the Met at 143, and relatively positive or more hydrophilic downstream. In comparison,

T123N and M133I/V/T missed in diagnostic assays presented more altered antigenic index profiles, while T143L showed similar if not lesser degree of changes (Table 1).

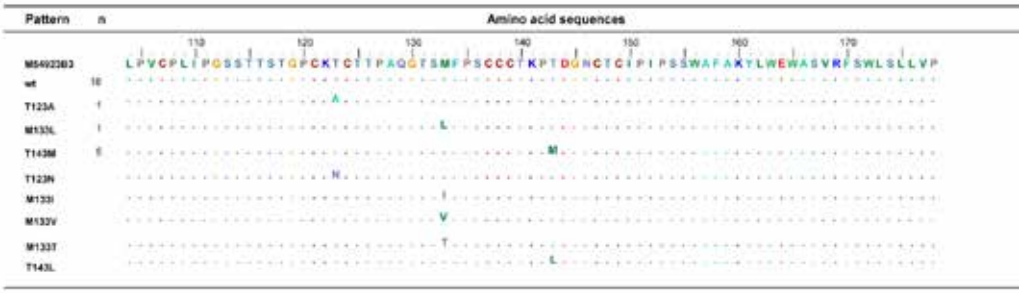


Figure 1. Alignment of amino acid sequences of HBV isolates in Indonesian blood donors with frequently-reported variants associated with failure of diagnostic assays. Three amino acid substitutions were identified in 7 HBV isolates in blood donors: Pattern 1, T123A, in one isolate; Pattern 2, M133L, in one isolate; Pattern 3, T143M, in five isolates. HBV DNA isolated from the remaining 18 samples showed wild type (wt) sequences with no amino acid substitution. Consensus of each of the three single mutation patterns and wt were aligned with five known variants frequently associated with problems in diagnostic assays and/or escape to vaccine/HBIg therapy: T123N, M133I, M133T, M133V, and T143L, together with M54923 sequence (genotype B/adw) retrieved from GenBank as a reference.

Tertiary structure prediction

The tertiary structure prediction of each variant isolated from Indonesian blood donors differed slightly from the wild type reference sequence, particularly in the *'a'* determinant region (Fig.2). The structure of the mainframe, which consisted mainly of helical structures, tended to be retained in all sequences, while the loop structures, including the *'a'* determinant, tended to differ slightly between these sequences. In pattern T123A, the loop containing the *'a'* determinant seemed to shift slightly compared to the reference wild-type. Although the side chain of Ala did not differ much in its orientation and position, the remainder of the loop shifted noticeably, as could be seen in the difference of the coiling and bends of the loop that made the contour of the *'a'* determinant against the cavity in the mainframe helices. Similar shift in loop structure was observed in pattern M133L, as could be shown in the different orientation of Leu side chain in position 133 compared to Met side chain in the wild-type. The pattern T143M, on the other hand, besides showing differentially-oriented side chain of Met, also showed significant changes in larger part of the loop. Larger region of the loop N-terminally of position 143 seemed to uncoil, while the loop positioned C-terminally of residue 143 bent closer toward the mainframe cavity compared to the reference structure.

Table 1. The Jameson-Wolf antigenicity index prediction of HBsAg within amino acid 118 – 160

Position	M54923		Pattern T123A*		Pattern T123N**		Pattern M133L*		Pattern M133I**		Pattern M133T**		Pattern M133V**		Pattern T143M*		Pattern T143L**	
	Residue	Antigenic Index	Residue	Antigenic Index	Residue	Antigenic Index	Residue	Antigenic Index	Residue	Antigenic Index	Residue	Antigenic Index	Residue	Antigenic Index	Residue	Antigenic Index	Residue	Antigenic Index
110	Ile	-0.2	Ile	-0.2	Ile	-0.2	Ile	-0.2	Ile	-0.2	Ile	-0.2	Ile	-0.2	Ile	-0.2	Ile	-0.2
111	Pro	-0.05	Pro	-0.05	Pro	-0.05	Pro	-0.05	Pro	-0.05	Pro	-0.05	Pro	-0.05	Pro	-0.05	Pro	-0.05
112	Gly	0.35	Gly	0.35	Gly	0.35	Gly	0.35	Gly	0.35	Gly	0.35	Gly	0.35	Gly	0.35	Gly	0.35
113	Ser	0.8	Ser	0.8	Ser	0.8	Ser	0.8	Ser	0.8	Ser	0.8	Ser	0.8	Ser	0.8	Ser	0.8
114	Ser	1	Ser	1	Ser	1	Ser	1	Ser	1	Ser	1	Ser	1	Ser	1	Ser	1
115	Thr	0.4	Thr	0.4	Thr	0.4	Thr	0.4	Thr	0.4	Thr	0.4	Thr	0.4	Thr	0.4	Thr	0.4
116	Thr	1.05	Thr	1.05	Thr	***0.8***	Thr	1.05	Thr	1.05	Thr	1.05	Thr	1.05	Thr	1.05	Thr	1.05
117	Ser	1.3	Ser	1.3	Ser	***1.05***	Ser	1.3	Ser	1.3	Ser	1.3	Ser	1.3	Ser	1.3	Ser	1.3
118	Thr	1.2	Thr	1.2	Thr	***0.95***	Thr	1.2	Thr	1.2	Thr	1.2	Thr	1.2	Thr	1.2	Thr	1.2
119	Gly	2.05	Gly	2.05	Gly	***1.6***	Gly	2.05	Gly	2.05	Gly	2.05	Gly	2.05	Gly	2.05	Gly	2.05
120	Pro	2.5	Pro	2.5	Pro	***2.05***	Pro	2.5	Pro	2.5	Pro	2.5	Pro	2.5	Pro	2.5	Pro	2.5
121	Cys	2.25	Cys	2.25	Cys	***2.5***	Cys	2.25	Cys	2.25	Cys	2.25	Cys	2.25	Cys	2.25	Cys	2.25
122	Lys	2	Lys	***1.6***	Lys	***2.25***	Lys	2	Lys	2	Lys	2	Lys	2	Lys	2	Lys	2
123	Thr	0.35	**Ala**	***0.4***	**Asn**	***2***	Thr	0.35	Thr	0.35	Thr	0.35	Thr	0.35	Thr	0.35	Thr	0.35
124	Cys	0.25	Cys	***0.15***	Cys	***1.5***	Cys	0.25	Cys	0.25	Cys	0.25	Cys	0.25	Cys	0.25	Cys	0.25
125	Thr	0.45	Thr	***0.65***	Thr	***0.9***	Thr	0.45	Thr	0.45	Thr	0.45	Thr	0.45	Thr	0.45	Thr	0.45
126	Thr	0.25	Thr	0.25	Thr	0.25	Thr	0.25	Thr	0.25	Thr	0.25	Thr	0.25	Thr	0.25	Thr	0.25
127	Pro	0.4	Pro	0.4	Pro	0.4	Pro	0.4	Pro	0.4	Pro	0.4	Pro	0.4	Pro	0.4	Pro	0.4
128	Ala	0.8	Ala	0.8	Ala	0.8	Ala	0.8	Ala	0.8	Ala	0.8	Ala	0.8	Ala	0.8	Ala	0.8
129	Gln	0.8	Gln	0.8	Gln	***0.4***	Gln	0.8	Gln	0.8	Gln	0.8	Gln	0.8	Gln	0.8	Gln	0.8
130	Gly	0.65	Gly	0.65	Gly	0.65	Gly	0.65	Gly	***0.25***	Gly	***0.8***	Gly	***0.25***	Gly	0.65	Gly	0.65
131	Thr	0.35	Thr	0.35	Thr	0.35	Thr	0.35	Thr	***-0.05***	Thr	0.35	Thr	***-0.05***	Thr	0.35	Thr	0.35
132	Ser	0.35	Ser	0.35	Ser	0.35	Ser	0.35	Ser	***-0.05***	Ser	***0.65***	Ser	***-0.05***	Ser	0.35	Ser	0.35
133	Met	-0.2	Met	-0.2	Met	-0.2	**Leu**	***-0.05***	**Ile**	***-0.6***	**Thr**	***0.8***	**Val**	***-0.45***	Met	-0.2	Met	-0.2
134	Phe	-0.2	Phe	-0.2	Phe	-0.2	Phe	-0.2	Phe	-0.2	Phe	***-0.05***	Phe	-0.2	Phe	-0.2	Phe	-0.2
135	Pro	0.2	Pro	0.2	Pro	0.2	Pro	0.2	Pro	0.2	Pro	***0.35***	Pro	0.2	Pro	0.2	Pro	0.2
136	Ser	0.2	Ser	0.2	Ser	0.2	Ser	0.2	Ser	0.2	Ser	0.2	Ser	0.2	Ser	0.2	Ser	0.2
137	Cys	0.2	Cys	0.2	Cys	0.2	Cys	0.2	Cys	0.2	Cys	0.2	Cys	0.2	Cys	0.2	Cys	0.2
138	Cys	0.64	Cys	0.64	Cys	0.64	Cys	0.64	Cys	0.64	Cys	0.64	Cys	0.64	Cys	***0.3***	Cys	***0.3***
139	Cys	1.18	Cys	1.18	Cys	1.18	Cys	1.18	Cys	1.18	Cys	1.18	Cys	1.18	Cys	***0.5***	Cys	***0.8***
140	Thr	1.27	Thr	1.27	Thr	1.27	Thr	1.27	Thr	***1.67***	Thr	1.27	Thr	1.27	Thr	***0.41***	Thr	***0.7***

Table 1. continued

Position	M54923		Pattern T123A*		Pattern T123N**		Pattern M133L*		Pattern M133I**		Pattern M133T**		Pattern M133V**		Pattern T143M*		Pattern T143L**	
	Residue	Antigenic Index	Residue	Antigenic Index	Residue	Antigenic Index	Residue	Antigenic Index	Residue	Antigenic Index	Residue	Antigenic Index	Residue	Antigenic Index	Residue	Antigenic Index	Residue	Antigenic Index
141	Lys	2.36	Lys	2.36	Lys	2.36	Lys	2.36	Lys	2.36	Lys	2.36	Lys	2.36	Lys	***1.47***	Lys	***1.75***
142	Pro	3.4	Pro	3.4	Pro	3.4	Pro	3.4	Pro	3.4	Pro	3.4	Pro	3.4	Pro	***2.33***	Pro	***2.6***
143	Thr	2.86	Thr	2.86	Thr	2.86	Thr	2.86	Thr	2.86	Thr	2.86	Thr	2.86	**Met**	***2.74***	**Leu**	***3***
144	Asp	2.57	Asp	2.57	Asp	2.57	Asp	2.57	Asp	2.57	Asp	2.57	Asp	2.57	Asp	***3.1***	Asp	***2.45***
145	Gly	1.93	Gly	1.93	Gly	1.93	Gly	1.93	Gly	1.93	Gly	1.93	Gly	1.93	Gly	***2.49***	Gly	***2.15***
146	Asn	1.59	Asn	1.59	Asn	1.59	Asn	1.59	Asn	1.59	Asn	1.59	Asn	1.59	Asn	***1.58***	Asn	***1.25***
147	Cys	0.1	Cys	0.1	Cys	0.1	Cys	0.1	Cys	0.1	Cys	0.1	Cys	0.1	Cys	***0.72***	Cys	***0.4***
148	Thr	-0.6	Thr	-0.6	Thr	-0.6	Thr	-0.6	Thr	-0.6	Thr	-0.6	Thr	-0.6	Thr	***-0.29***	Thr	-0.6
149	Cys	-0.6	Cys	-0.6	Cys	-0.6	Cys	-0.6	Cys	-0.6	Cys	-0.6	Cys	-0.6	Cys	-0.6	Cys	-0.6
150	Ile	-0.6	Ile	-0.6	Ile	-0.6	Ile	-0.6	Ile	-0.6	Ile	-0.6	Ile	-0.6	Ile	-0.6	Ile	-0.6
151	Pro	-0.6	Pro	-0.6	Pro	-0.6	Pro	-0.6	Pro	-0.6	Pro	-0.6	Pro	-0.6	Pro	-0.6	Pro	-0.6
152	Ile	-0.45	Ile	-0.45	Ile	-0.45	Ile	-0.45	Ile	-0.45	Ile	-0.45	Ile	-0.45	Ile	-0.45	Ile	-0.45
153	Pro	-0.05	Pro	-0.05	Pro	-0.05	Pro	-0.05	Pro	-0.05	Pro	-0.05	Pro	-0.05	Pro	-0.05	Pro	-0.05
154	Ser	0.35	Ser	0.35	Ser	0.35	Ser	0.35	Ser	0.35	Ser	0.35	Ser	0.35	Ser	0.35	Ser	0.35
155	Ser	0.35	Ser	0.35	Ser	0.35	Ser	0.35	Ser	0.35	Ser	0.35	Ser	0.35	Ser	0.35	Ser	0.35
156	Trp	-0.2	Trp	-0.2	Trp	-0.2	Trp	-0.2	Trp	-0.2	Trp	-0.2	Trp	-0.2	Trp	-0.2	Trp	-0.2
157	Ala	-0.6	Ala	-0.6	Ala	-0.6	Ala	-0.6	Ala	-0.6	Ala	-0.6	Ala	-0.6	Ala	-0.6	Ala	-0.6
158	Phe	-0.6	Phe	-0.6	Phe	-0.6	Phe	-0.6	Phe	-0.6	Phe	-0.6	Phe	-0.6	Phe	-0.6	Phe	-0.6
159	Ala	-0.6	Ala	-0.6	Ala	-0.6	Ala	-0.6	Ala	-0.6	Ala	-0.6	Ala	-0.6	Ala	-0.6	Ala	-0.6
160	Lys	-0.6	Lys	-0.6	Lys	-0.6	Lys	-0.6	Lys	-0.6	Lys	-0.6	Lys	-0.6	Lys	-0.6	Lys	-0.6

*Variants found in Indonesian blood donor; **Variants frequently associated with problems in diagnostic assays and/or escape to vaccine/HBIg therapy. Residues with substitutions and their positions are shown in **bold**. Altered antigenicity index of affected residues in each substitution pattern are shown in **bold** and *italics*: T123A alters four consecutive residues (aa 122–125); M133L alters the antigenic index of position 133 only; T123N, M133I/T/V, and T143L cause relatively extensive antigenic index changes in 11, 5, 5, 4, and 10 residues, respectively; T143M shows the most significant changes in the antigenic profile of HBsAg between residues 138 to 148.

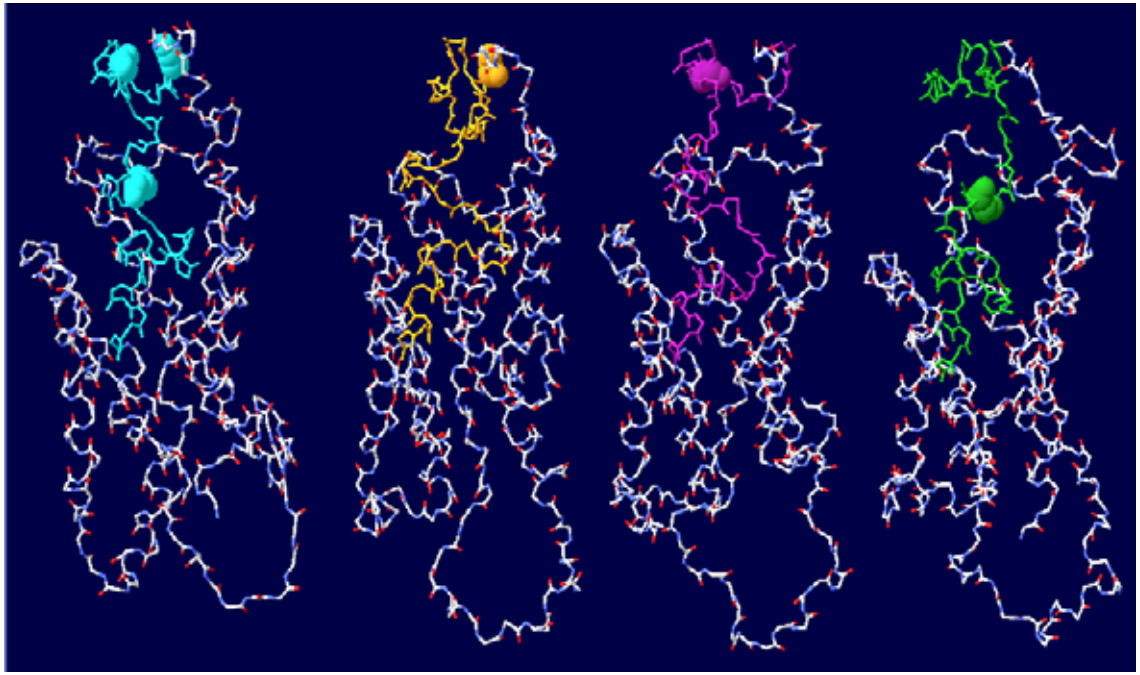


Figure 2. Comparison of tertiary structure prediction. Tertiary structure prediction of M54923 (reference sequence), T123A, M133L, and T143M mutants. The 'a' determinant is shown in blue, yellow, magenta, and green, respectively, while residues of importance are labelled with the side chains shown.

Discussion

HBV mechanism of replication includes an RNA intermediate that is reverse-transcribed into DNA by error-prone RNA polymerase [30]. This process results in a high mutation rate of approximately 1.4-3.2 x 10^{-5} substitutions/site/year for the whole genome and even higher for the surface gene [30-31]. This allows the virus to evolve within a chronically infected individual to form a naturally occurring quasi-species pool of HBV variants [5,29]. In regions with high HBV endemicity, the relatively high rate of viral transmission might provide more opportunities for super-infection and multiple infections to occur, which would result in increased number of variants circulating within individuals as well as in the population [2,32]. The composition of variants in the viral population is maintained by its environment. Variants better suited to the host environment would prevail and dominate the population [33]. In such cases, environmental changes induced by either natural immune response, vaccine-induced or therapeutic immunoglobulin (HBIg), or even anti-viral therapy may select for variants that can evade these protective measures, particularly those exhibiting mutation-induced conformational changes at the antigenic *'a'* determinant of its surface antigen [2-3,5]. Selection of variants is usually indicated by certain serological markers, such as isolated anti-HBc, co-occurrence of both HBsAg and anti-HBs, and inconsistent HBsAg assay results [34]. The presence of these variants poses potential threat to the success of vaccination and supply of safe blood products due to the possible evasion from vaccine-generated antibody and poor detection by the available diagnostic assays [6].

Numerous studies have shown that three dimensional conformations of proteins contribute toward their biological functions as well as their interactions with other molecules [35-36]. Substitutions of key amino acid residues may affect the stability and structure of a protein, altering its properties and interactions with other particles. Protein modelling of HBsAg variants might give insight into the structural basis of HBV variation at the molecular level, and how it affects the HBsAg recognition by its specific antibody.

Substitutions of Thr 123, Met 133 and Thr 143 into other amino acid residues as found in this study had been described in relation to failure of HBIg therapy and problems in detection assays [5,24-29,37-38]. The outcome of these substitutions is related to the site of mutation and the property of the respective amino acid, which is also observed in the mutants found in this study. Thr123, although located upstream of the *'a'* determinant, is in close proximity to the Cys 124 residue responsible for maintaining the integrity of HBsAg antigenic loop. There had been reports of insertions between Cys residues 121 and 124 that reduced or abolished bindings by monoclonal antibodies [39-40]. Furthermore, in a study by Chen *et al.*, the preservation of Thr at residue 123

seemed to be an important factor in the recognition of one of the *'a'* determinant epitopes by monoclonal anti-HBs [7]. Hence, the substitution site is important because it may disturb the disulphide bonds, leading to the alteration of loop conformation and decrease or loss of neutralizing antibody binding.

The other two mutation sites, Met 133 and Thr 143, are located within the first (aa 124-137) and second (aa 139-147) antigenic loops of the *'a'* determinant, respectively [7-8,41]. Ample reports on substitutions within these two regions had been published [5,24,26-27,37,40-43], as the *'a'* determinant is known as the main antibody recognition site of HBsAg. Mutations at these regions would predictably affect the loop conformation and causes problems of escape mutants and diagnostic failure.

As of the property of each amino acid, protein is a macromolecule made of monomeric amino acids. Each amino acid has distinct properties attributable to its side-chain, and the structure of a protein is dependent on the composition of its amino acids [44]. Therefore, differences in amino acid properties might contribute to the changes in the structure of the *'a'* determinant loop. Methionine, Alanine, Leucine, Isoleucine, and Valine are amino acids with non-polar, aliphatic side chains, while Threonine and Aspargine have a polar although uncharged side chain (-CH(CH_3)-OH and -CH_2-CO-NH_2 groups). Within the non-polar, aliphatic amino acids themselves, there are differences in the length and bulkiness of the side chain; alanine has a methyl group (-CH_3), valine with iso-propyl group (-CH(CH_3)-CH_3), leucine with iso-butyl group (-CH_2-CH(CH_3)-CH_3), isoleucine with 2º-butyl group (-CH(CH_3)- CH_2-CH_3) and methionine with a methyl-ethyl-sulphide group (–CH_2-CH_2-S-CH_3). These slight differences in the amino acid properties may affect the tertiary structure of the protein, as different polarity determines the hydrophobicity of the residue, while differences in length and bulkiness of the side chain may influence the steric hindrance between neighbouring residues [44].

The degree of changes in antigenicity profile was highest in T143M pattern, followed closely by T123N and T143L, then lesser changes in M133I/T/V as well as T123A and M133L. M133L mutant showed the least significant changes, probably because it is located in less-antigenic first loop [41], and also because both Met and Leu are non-polar residues with similar bulkiness of their side-chains. T123A mutant, on the other hand, involved changes from a polar Thr into a non-polar and slightly smaller Ala. Although it may affect the conformation by means of influencing the disulphide bond, the effect would be minimized because of the nature and size of Ala. The trend in M133I/T/V can also be correlated with the differential amino acid properties, with similar changes between M133I and M133V that involve similarly-sized non polar Met, Ile, and Val; and slightly more significant antigenic alteration in M133T, in which there is a change from

Met to polar Thr. Marked changes were also observed in T123N and T143L substitutions, which might be caused by both the shift from slightly small, polar Thr into either larger, more polar Asp or bulkier, non-polar Leu and the importance of their respective locations. Similarly, in T143M mutation, a major change from polar Thr into non-polar, significantly bulkier Met within the more antigenic second loop of the *'a'* determinant occurred [41]. This is also seen when several of the substitution patterns were constructed in tertiary structure modelling (Fig.2), with more significant changes observed if the amino acids involved had higher degree of variation in their properties.

Comparison of variants T123A, M133L and T143M with the reference wild-type HBsAg showed different predicted tertiary structures with lesser degree of changes observed in the mainframe helices compared to the loops' structures (Fig.2). This might be caused by the higher degree of freedom in the movement of the loop regions. Loop regions tend to be hydrophilic and interact more freely with the surrounding environment, while mainframe helices are much more constrained in structure due to the hydrophobicity and tendency to maintain the distance between their residues [44].

All these observations were obtained by mathematical model and prediction software, involving various algorithms to calculate the antigenic index and methods to predict variant HBsAg conformation. Further analysis involving experimental studies of the interaction between variant HBsAg and anti-HBs is needed to confirm these preliminary findings, and continuous screening of larger sets of samples is necessary to obtain more data on the emergence of new variants that might circulate in the population.

Conclusions

In conclusion, antigenic index analysis and *de novo* prediction of tertiary conformation of the three HBsAg variants (T123A, M133L, and T143M) found in Indonesian blood donor samples with undetectable HBsAg revealed that T143M substitution altered the antigenicity most significantly compared to the other two mutation patterns and the other known variants. This finding offers insight into the possibility of predicting antigenic changes in unique variants based on its primary amino acid sequence. It also underlines the importance of protein structure prediction in understanding the dynamic interactions between pathogenic agents and host immune system, in anticipation of new variants that might emerge in the future. This would in turn be a useful tool to better overcome the issues regarding detection failure by diagnostic assays and the global use of vaccines, particularly in endemic areas, as one possible mechanism of selecting escape mutants.

Acknowledgements

The authors would like to express their gratitude to the Indonesian Blood Transfusion Units in Medan and Solo, Professor J. Tarigan from the Faculty of Medicine, North Sumatra University, Medan, and Professor F.X. Suparyatmo from the Faculty of Medicine, University of Sebelas Maret, Solo, Indonesia, for their donation of blood donors samples.

Authors' contributions

SII carried out the protein prediction analysis, participated in the sequence alignment and drafted the manuscript. MDT carried out the molecular genetic studies, sequence analysis, and the design of the study. MR participated in the serological and molecular genetic studies. DHM conceived of the study, and participated in its design and coordination and helped to draft the manuscript. All authors read and approved the final manuscript.

Competing interests

The author(s) declare that they have no competing interests.

References

1. Kann M: **Structural and Molecular Virology.** In: *Hepatitis B Virus*. Edited by Lai CL, Locarnini S. London, International Medical Press Ltd. 2002, 9-22.
2. Weber B: **Diagnostic Impact of the Genetic Variability of the Hepatitis B Virus Surface Antigen Gene.** *J Med Virol* 2006, **78**:S59-S65.
3. Waters JA, Kennedy M, Voet P, Hauser P, Petre J, Carman W, Thomas HC: **Loss of the Common "A" Determinant of Hepatitis B Surface Antigen by a Vaccine-Induced Escape Mutant.** *J Clin Invest* 1992, **90**:2543-2547.
4. Carman WF, Trautwein C, van Deursen FJ, Colman K, Dornan E, McIntyre G, Waters J, Kliem V, Muller R, Thomas HC, Manns MP: **Hepatitis B Virus Envelope Variation After Transplantation With and Without Hepatitis B Immune Globulin Prophylaxis.** *Hepatology* 1996, **24**:489-493.
5. Carman WF, van Deursen FJ, Mimms LT, Hardie D, Coppola R, Decker R, Sanders R: **The Prevalence of Surface Antigen Variants of Hepatitis B Virus in Papua New Guinea, South Africa, and Sardinia.** *Hepatology* 1997, **26**:1658-1666.

6. Carman WF, Zanetti AR, Karayiannis P, Waters J, Manzillo G, Tanzi E, Zuckerman AJ, Thomas HC: **Vaccine-induced escape mutant of hepatitis B virus.** *Lancet* 1990, **336**:325-329.

7. Chen YCJ, Delbrook K, Dealwis C, Mimms L, Mushahwar IK: **Discontinuous epitopes of hepatitis B surface antigen derived from a filamentous phage peptide library.** *Proc Natl Acad Sci USA* 1996, **93**:1997-2001.

8. Stirk HJ, Thornton JM, Howard CR: **A Topological Model for Hepatitis B Surface Antigen.** *Intervirology* 1992, **33**:148-158.

9. Thedja MD, Roni M, Harahap AR, Siregar NC, Ie SI, Muljono DH: **Occult hepatitis B in blood donors in Indonesia: altered antigenicity of the hepatitis B virus surface protein.** *Hepatol Int* 2010, **4**:608–614.

10. Raimondo G, Allain JP, Brunetto MR, Buendia MA, Chen DS, Colombo M, Craxi A, Donato F, Ferrari C, Gaeta GB, Gerlich WH, Levrero M, Locarnini S, Michalak T, Mondelli MU, Pawlotsky JM, Pollicino T, Prati D, Puoti M, Samuel D, Shouval D, Smedile A, Squadrito G, Trepo C, Villa E, Will H, Zanetti AR, Zoulim F: **Statements from the Taormina expert meeting on occult hepatitis B virus infection.** *J Hepatol* 2008, **49**:652–657.

11. Sastrosoewignjo RI, Omi S, Okamoto H, Mayumi M, Rustam M, Sujudi T: **The complete nucleotide sequence of HBV DNA clone of subtype adw (pMND122) from Menado in Sulawesi Island, Indonesia.** *ICMR Ann* 1987, **7**:51-60.

12. Hall TA: **BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT.** *Nucl Acids Symp Ser* 1999, **41**:95-98.

13. Jameson BA, Wolf H: **The antigenic index: a novel algorithm for predicting antigenic determinants.** *Cabios* 1988, **4(1)**:181-186.

14. Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W, Lipman DJ: **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.** *Nucleic Acids Res* 1997, **25**:3389-3402.

15. Zdobnov EM, Apweiler R: **InterProScan - an integration platform for the signature-recognition methods in InterPro.** *Bioinformatics* 2001, **17**:847-848.

16. Mulder NJ, Apweiler R, Attwood TK, Bairoch A, Barrell D, Bateman A, Binns D, Biswas M, Bradley P, Bork P, Bucher P, Copley RR, Courcelle E, Das U, Durbin R, Falquet L, Fleischmann W, Griffiths-Jones S, Haft D, Harte N, Hulo N, Kahn D, Kanapin A, Krestyaninova M, Lopez R, Letunic I, Lonsdale D, Silventoinen V, Orchard SE, Pagni M, Peyruc D, Ponting CP, Selengut JD, Servant F, Sigrist CJA, Vaughan R, Zdobnov EM: **The InterPro Database, 2003 brings increased coverage and new features.** *Nucleic Acids Res* 2003, **31**:315-318.

17. Söding J, Biegert A, Lupas AN: **The HHpred interactive server for protein homology detection and structure prediction.** *Nucleic Acids Res* 2005, **33**:W244-W248.

18. Westbrook J, Feng Z, Chen L, Yang H, Berman HM: **The Protein Data Bank and structural genomics.** *Nucleic Acids Res* 2003, **31**:489-491.

19. Zhang Y: **Progress and challenges in protein structure prediction.** *Current Opinion in Structural Biology* 2008, **18**:342-348.

20. Wu ST, Skolnick J, Zhang Y: **Ab initio modeling of small proteins by iterative TASSER simulations.** *BMC Biology* 2007, **5**:17.

21. Zhang Y: **Template-based modeling and free modeling by I-TASSER in CASP7.** *Proteins* 2007, **69(S8)**:108-117.

22. Zhang Y: **I-TASSER server for protein 3D structure prediction.** *BMC Bioinformatics* 2008, **9**:40-47.

23. Guex N, Peitsch MC: **SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling.** *Electrophoresis* 1997, **18**:2714-2723.

24. Wallace LA, Echevarria JE, Echevarria JM, Carman WF: **Molecular Characterization of Envelope Antigenic Variants of Hepatitis B Virus from Spain.** *J Infect Dis* 1994, **170**:1300-1303.

25. Carman WF: **The clinical significance of surface antigen variants of hepatitis B virus.** *J Viral Hepat* 1997, **4(Suppl.1)**:11-20.

26. Ireland JH, O'Donnell B, Basuni AA, Kean JD, Wallace LA, Lau GKK, Carman WF: **Reactivity of 13 *In Vitro* Expressed Hepatitis B Surface Antigen Variants in 7 Commercial Diagnostic Assays.** *Hepatology* 2000, **31**:1176-1182.

27. Hou JL, Wang ZH, Cheng JJ, Lin YL, Lau GKK, Sun J, Zhou FY, Waters J, Karayiannis P, Luo KX: **Prevalence of Naturally Occuring Surface Gene Variants of Hepatitis B Virus in Nonimmunized Surface Antigen-Negative Chinese Carriers.** *Hepatology* 2001, **34**:1027-1034.

28. Chen HB, Fang DX, Li FQ, Jing HY, Tan WG, Li SQ: **A novel hepatitits B virus mutant with A-to-G at nt551 in the surface antigen gene.** *World H Gastroenterol* 2003, **9(2)**:304-308.

29. Echevarria JM, Avellon A: **Hepatitis B Virus Genetic Diversity.** *J Med Virol* 2006, **78**:S36–S42.

30. Okamoto H, Imai M, Kametani M, Nakamura T, Mayumi M: **Genomic Heterogeneity of Hepatitis B Virus in a 54-Year-Old Woman Who Contracted the Infection through Materno-Fetal Transmission.** *J Exp Med* 1987, **57(4)**:231-236.

31. Zaaijer HL, Bouter S, Boot HJ: **Substitution rate of hepatitis B surface gene.** *J Viral Hepat* 2008, **15**:239-245.

32. Kramvis A, Kew M, Francois G: **Hepatitis B virus genotypes.** *Vaccine* 2005, 23:2409-2423.

33. Carman W, Thomas H, Domingo E: **Viral genetic variation: hepatitis B virus as clinical example.** *Lancet* 1993, **341**:349-353.

34. Gerlich WH: **Diagnostic Problems Caused by HBsAg Mutants – A Consensus Report of an Expert Meeting.** *Intervirology* 2004, **47**:310-313.

35. Fairchild S, Pachter R, Perrin R: **Protein Structure Analysis and Prediction.** *The Mathematica Journal* 1995, 5(4):64-69.

36. Ponsel D, Bruss V: **Mapping of Amino Acid Side Chains on the Surface of Hepatitis B Virus Capsids Required for the Envelopment and Virion Formation.** *J Virol* 2003, 77(1):416-422.

37. Levicnic-Stezinar S: Hepatitis B surface antigen escape mutant in a first time blood donor potentially missed by a routine screening assay. *Clin Lab* 2004, 50:49-51.

38. Ren FR, Tsubota A, Hirokawa T, Kumada H, Yang ZH, Tanaka H: **A unique amino acid substitution, T126I, in human genotype C of hepatitis B virus S gene and its possible influence on antigenic structural change.** *Gene* 2006, 383:43-51.

39. Carman WF, Korula J, Wallace L, MacPhee R, Mimms L, Decker R: **Fulminant reactivation of hepatitis B due to envelope protein mutant that escaped detection by monoclonal HBsAg ELISA.** *Lancet* 1995, **345**:1406-1407.

40. Seddigh-Tonekaboni S, Waters JA, Jeffers S, Gehrke R, Ofenloch B, Horsch A, Hess G, Thomas HC, Karayiannis P: **Effect of Variation in the Common "a" Determinant on the Antigenicity of Hepatitis B Surface Antigen.** *J Med Virol* 2000, **60**:113-121.

41. Zuckerman AJ: **Effect of hepatitis B virus mutants on efficacy of vaccination.** *Lancet* 2000, **355**:1382-1384.

42. Oon CJ, Lim GK, Ye Z, Goh KT, Tan KL, Yo SL, Hopes E, Harrisonn TJ, Zuckerman AJ: **Molecular epidemiology of hepatitis B virus vaccine variants in Singapore.** *Vaccine* 1995, **13(8)**:699–702.

43. Protzer-Knolle U, Naumann U, Bartenschlager R, Berg T, Hopf U, zum Buschenfelde KHM, Neuhaus P, Gerken G: **Hepatitis B Virus With Antigenically Altered Hepatitis B Surface Antigen Is Selected by High-Dose Hepatitis B Immune Globulin After Liver Transplantation.** *Hepatology* 1998, **27**:254-263.

44. Lehninger AL, Nelson DL, Cox MM: **Amino Acids and Peptides.** In *Lehninger: Principles of Biochemistry.* 2nd edition. New York: Worth Publishers; 1993:111-133.

CHAPTER 6

Viral Kinetics in the Natural History of Chronic Hepatitis B in Indonesia

Turyadi[1], Meta Dewi Thedja[1,2], Susan Irawati Ie[1], Alida R. Harahap[1,3], Korri El Khobar[1], Martono Roni[1], David Handojo Muljono[1,4,5]

[1]Eijkman Institute for Molecular Biology, Jakarta, Indonesia

[2]Eijkman Winkler Institute, Utrecht Medical Centre, Utrecht, Netherlands

[3]Clinical Pathology Department, Faculty of Medicine, University of Indonesia, Jakarta, Indonesia

[4]Faculty of Medicine, Hasanuddin University, Makassar, Indonesia

[5]Sydney Medical School, University of Sydney, Australia

In progress

Abstract

Introduction. Chronic hepatitis B (CHB) is a state of dynamic and complex interactions between hepatitis B virus (HBV) and host immunity. Understanding of natural course of CHB is important in patient management. We studied changes in viral markers by quantification of hepatitis B surface antigen (HBsAg), hepatitis B e antigen (HBeAg), and HBV-DNA levels, with the implication of HBV genotype, subtype, basal core promoter (BCP) T1762/A1764 and precore A1896 mutations.

Methods. One-hundred and fifty-two treatment-naïve CHB patients were classified into four phases: immune-tolerant (IT), immune-clearance (IC), low/non-replicative (LR), and 'e' negative hepatitis B (ENH), based on HBeAg status, ALT and HBV-DNA levels. HBV-DNA was detected and quantified by polymerase chain reaction then analyzed by sequencing. HBsAg and HBeAg levels were determined by serological assays.

Results. HBsAg and HBV-DNA levels varied between different CHB phases. HBsAg median level was highest in IT (4.22 $\log_{10}$IU/mL) and lowest in LR (2.52 $\log_{10}$IU/mL), while HBV-DNA median levels were high in IT and IC (4.55 $\log_{10}$IU/mL and 5.13 $\log_{10}$IU/mL) and lowest in LR (1.73 $\log_{10}$IU/mL). Increased levels of both markers were also seen in ENH. A significant correlation between HBsAg and HBV-DNA levels was observed in IT and IC ($r=0.712$; $p<0.001$ and $r=0.731$; $p<0.001$), modest in ENH, but missing in LR. No correlation was observed between HBeAg and HBV-DNA. The samples were dominated by HBV genotype B (77.2%) and C (22.8%), with subtypes *adw* (75.2%), *adr* (22%), and *ayw* (2.8%). Genotype/subtype distribution was comparable in all phases. Precore mutants was more prevalent in HBeAg-negative group ($p=0.004$), while BCP mutations occurred similarly in all phases.

Conclusion. HBsAg and HBV DNA levels were high and strongly correlated in early CHB phases. The correlation dissociated after HBeAg seroconversion, indicating different immune controls affecting HBV replication and HBsAg production. This correlation was later documented in ENH, supporting the view that ENH is a variant of IC. Comparable distribution of BCP mutations signify that the deleterious impact of these mutations could occur at any time during the disease course. A temporal association between HBeAg seroconversion and the increase of the prevalence of BCP/precore mutations suggests the role of immune pressure in driving HBV evolution. This study calls for further evaluation of quantitative HBsAg and HBeAg assay performance in clinical practice before including them in treatment algorithm. Pre- and on-treatment studies – stratified based on host ethnic background and viral genetic characteristics – are necessary.

Introduction

Hepatitis B virus (HBV) infection has a global distribution with an estimated of more than 2 billion people infected worldwide, and approximately 240 million suffering from chronic infection with risks of developing serious illnesses such as cirrhosis and hepatocellular carcinoma (HCC).[1] The majority of patient with chronic HBV infection reside in Asia including Indonesia, a country with moderate-to-high hepatitis B endemicity.[2]

Chronic Hepatitis B (CHB) is determined by seropositivity of hepatitis B surface antigen (HBsAg) of more than six months.[3] Persistent viral infection in CHB is achieved by the establishment and maintenance of a stable pool of covalently closed circular DNA (cccDNA) in the nuclei of infected hepatocytes.[4] The cccDNA acts as a template for the transcription of all viral genes necessary for protein production and viral replication.[5] The natural course of CHB can be divided into several phases based on the viral-host interaction: immune tolerance (IT), immune clearance (IC), and low/non replicative (LR).[6] In IT phase, viral replication proceeds at a very high level and hepatitis B "e" antigen (HBeAg) is detectable in serum, while in IC phase, HBV DNA levels tend to fluctuate and will decrease progressively. HBeAg can still be detected until clearance by the immune system at the end of IC phase. The third phase, LR, is characterized by HBeAg seroconversion and undetectable or low level of HBV DNA, followed by sustained clinical remission (inactive carrier stage). Most patients remain in this phase for many years, if not indefinitely.[7] However, this inactive state can revert back to IC phase with HBeAg seropositive, or maintain HBeAg seronegative and develop into HBeAg-negative hepatitis (ENH)[8], which is also categorized as a separate phase of the natural history of CHB.[9] ENH phase is characterized by elevation of HBV replication with undetectable HBeAg.[10] The phases of the natural history of CHB are not necessarily present sequentially and not all patients go through every phase.[9,11]

During the course of infection, naturally occurring HBV mutants may present under the pressure of host immunity.[12] The most critical mutations with implication on the natural course of CHB are the A1762T/G1764A double mutations in the basal core promoter (BCP) region and G1896A mutation in the precore region, hereafter referred to as BCP and precore mutations, that diminish the production of HBeAg during HBV replication.[13] Advanced technique development has allowed for the quantitative assessment of serum HBV DNA levels as well as HBsAg levels that have strong correlation with cccDNA as the intrahepatic template for viral replication.[14,15,16]

Understanding of the natural history of CHB is absolutely needed to help clinicians in management of CHB patients. However, the relation between viral factors including HBsAg, HBeAg and HBV DNA levels, HBV genotype, and BCP and precore mutations in the natural history of CHB have not been well characterized, especially in Indonesia. This study was carried out with aims to study the changes of viral markers in CHB patients in Indonesia, measured by quantitative HBsAg and HBeAg levels that reflect the transcriptional activity, and HBV DNA level as the parameter of viral replication. Further, the implication of BCP and precore mutations on the natural history of the disease was also investigated.

Materials and Methods

Patients

A total of 152 treatment-naïve patients positive for HBsAg for at least 6 months were referred to the Eijkman Institute for Molecular Biology, Jakarta, and enrolled in this study following written consent from the patients with agreement from the referring physicians. The study was approved by Eijkman Institute Research Ethics Commission (EIREC No. 23/2007). Patients co-infected with hepatitis C virus were excluded from this study. HBV DNA levels were tested soon after sera collection, while the remaining sera were stored at -80 ^{o}C and later used for HBsAg and HBeAg levels quantitation, viral genotyping, subtyping, and analysis of BCP and precore mutations. The patients comprised 102 males and 50 females (14–80 years old). All patients were categorized into the four phases of CHB based on their HBeAg serostatus, HBV viral load and serum ALT level. IT phase was defined as HBeAg seropositive, HBV DNA $\geq 10^4$ IU/mL and serum ALT <2X upper limit normal (ULN). IC phase was defined as HBeAg seropositive and serum ALT >2X ULN. LR phase was defined as HBeAg seronegative, ALT <2X ULN and HBV DNA $<10^4$ IU/mL, whereas ENH was defined as HBeAg seronegative and ALT >2X ULN.[6,17,18,19]

HBV Serological Assays

HBeAg and anti-HBe were tested using Monolisa^{TM} HBeAg Ag-Ab PLUS (Biorad,France) enzyme immunoassay kit according to the manufacturer's instructions for qualitative determination of HBeAg and anti-HBe. HBeAg-positive samples underwent quantitative measurement by Elecsys® HBeAg assay (Roche Diagnostic GmbH, Manheim, Germany) based on the manufacturer's instruction. Serum HBsAg was quantified using

ARCHITECT Plus i2000SR (Abbott Laboratories, Chicago, US), with dynamic range of 0.05–250 IU/mL, according to the manufacturer's instructions. Samples with HBsAg levels higher than 250 IU/mL were diluted 150 or 500 times using the ARCHITECT HBsAg Manual Diluent (Abbott Laboratories, Chicago USA) prior to measurement.

HBV DNA Analysis

HBV-DNA titer was determined from 500 µL of serum using quantitative realtime PCR (Cobas Taqman^{TM} HBV Test, Roche Diagnostics, Indianapolis, USA), according to the manufacturer's instructions, with range of linearity between 6 - 1.1x10^{8} IU/mL. HBV DNA was extracted from 140 µL of serum using QIAamp DNA MiniKit (QIAGEN, CA, USA) according to the manufacturer's instructions. The extracted DNA was eluted in 60 µL of elution buffer. DNA fragment of the 'a determinant' region of S gene was amplified by nested polymerase chain reaction (PCR) using specific primer sets S2-1 (nt 455-474; 5'-CAA GGT ATG TTG CCC GTT TG-3') and S1-2 (nt 704-685; 5'-CGA ACC ACT GAA CAA ATG GC-3') as outer primers[20] and S88 (nt 462–481; 5'-TGT TGC CCG TTT GTC CTC TA-3')[21] and S2-2 (nt 687-668; 5'-GGC ACT AGT AAA CTG AGC CA-3') as internal primers.[20] Positive amplification products were purified using PCR purification column (QIAGEN, CA, USA) and subjected to direct sequencing reaction on DNA sequence analyzer ABI 3130xl (Applied Biosystems, USA).

HBV genotype was determined by phylogenetic analysis based on the 226-nucletide-sequences of surface gene compared with 70 reference sequences of known genotype (A-H) retrieved from the Genbank, using Phylip 3.68 software with Kimura-2 parameter, neighbor-joining algorithm, and 1000 bootstrapping.[22] HBV subtype was determined based on deduced amino acids at positions 122 and 160 of the HBsAg as previously described.[23,24]

Determination of BCP and Precore Mutants

Amplification of BCP and precore regions was done by nested PCR using PC1 (nt 1554-1573; 5'- CTG TGC CTT CTC ATC TGC CG-3') and PC2 (nt 1972–1949; 5'- AAA GAA GTC AGA AGG CAA AAA AGA-3') as outer primers and S12 (nt 1679 – 1699; 5'-AAT GTC ACC GAC CGA CCT TG-3') and S13 (nt 1941-1919; 5'-TCC ACA GAA GCT CCA AAT TCT AA-3') as internal primers. Positive amplification products were purified and sequenced as described previously. The sequences were aligned with wild type reference sequence (Accession number M54923) retrieved from the GenBank. All nucleotide numberings were based on *Eco*RI restriction site within the HBV genome.

Statistical analysis

Continuous and categorical variables were compared between groups using the Mann-Whitney test and the chi-square/Fisher's exact test, respectively. Overall comparison was analyzed by ANOVA. Pearson's correlation coefficient (r) was used to describe the correlation between two continuous, normally distributed variables. Spearman's correlation was used where variables were not normally distributed. All statistical analyses were performed using Statistical Package for Social Sciences (SPSS) v.16 (SPSS Inc., Chicago, IL). All statistical significance values were assessed at $p < 0.05$.

Results

Study population

Of 152 CHB patients enrolled in this cross-sectional study, 65 (42.8%) were HBeAg-positive, while the remaining 87 (57.2%) were HBeAg-negative. Patients were classified into the four phases of CHB: IT (33; 21.7%), IC (32; 21.1%), LR (34; 22.4%) and ENH (53; 34.9%). HBeAg-positive (IT and IC) patients were significantly younger than HBeAg-negative (LR and ENH) patients (p=0.002). Gender distribution was comparable between IT, IC and LR phases, but there was a significant male dominance in ENH group (p=0.006). Detailed characteristics of the study population are shown in Table 1.

Clinical characteristics according to chronic hepatitis B phases

AST and ALT levels varied across the CHB phases, and were significantly higher among the IC and ENH patients compared to the IT and LR. The median AST/ALT levels in IT, IC, LR, and ENH were 29/36 U/L, 129.5/126 U/L, 27/31.5 U/L, and 73/87 U/L, respectively.

HBsAg, HBeAg and HBV DNA levels in chronic hepatitis B phases

As seen in Table 1, the HBsAg and HBV DNA levels varied between patients from different CHB phases (p=0.001 and p <0.001, respectively). The HBsAg median level was highest in IT (4.22 $\log_{10}$ IU/mL), decreased in IC (3.34 $\log_{10}$ IU/mL), lowest in LR (2.52 $\log_{10}$ IU/mL), but later increased in ENH (3.37 $\log_{10}$ IU/mL). The HBV DNA median levels were high in IT and IC phases (4.55 $\log_{10}$ IU/mL and 5.13 $\log_{10}$ IU/mL, respectively), lowest in LR (1.73 $\log_{10}$ IU/mL), then increased again (4.83 $\log_{10}$ IU/mL) in ENH. The HBsAg-to-HBV DNA ratio was significantly higher in the LR phase compared

to IT, IC and ENH (1.1 vs 0.86, 0.60, and 0.64; p=0.029). Quantitative HBeAg measurement was done for HBeAg-positive patients, with comparable median values of HBeAg levels at 592.0 PEIU/mL and 236.6 PEIU/mL in the IT and IC phases, respectively (p=0.803).

Overall and individual correlations between HBsAg, HBeAg, HBV DNA levels, ALT levels, and age among CHB phases

Analysis using Spearman correlation test showed significant overall correlation between HBsAg and HBV DNA levels (r=0.659; p<0.001). Stronger correlation was observed in patients with positive HBeAg compared to those with negative HBeAg (r=0.661; p<0.001 vs r=0.586; p<0.001). In the HBeAg-positive group, HBsAg levels also had positive correlation with HBeAg levels (r=0.481; p=0.005), while there was no correlation between HBeAg and HBV DNA levels (r=0.133; p=0.460). As shown in Fig. 1 and Table 2, analysis for individual phases revealed that the correlations between HBsAg and HBV DNA levels were strong in IT and IC (r=0.712; p<0.001 and r=0.731; p=<0.001, respectively), moderate in ENH (r=0.429; p=0.002), but insignificant in LR (r=0.288; p=0.098).

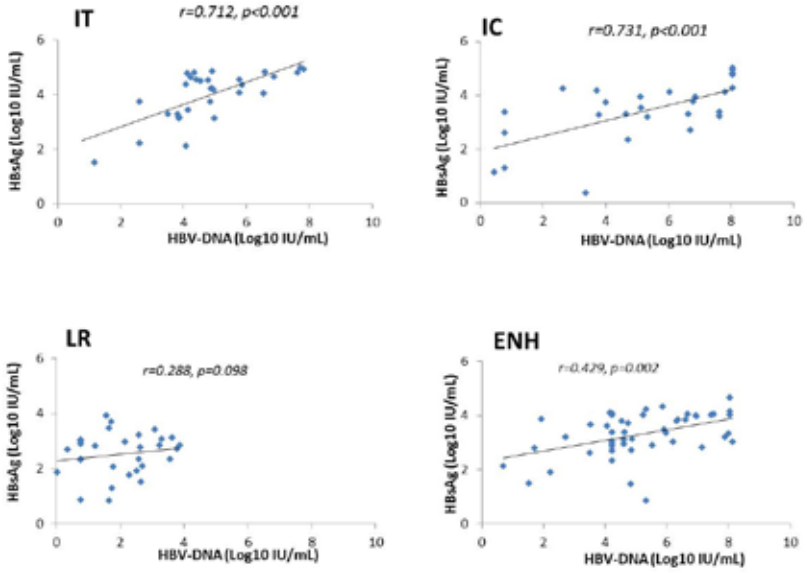


Figure 1. Distribution and correlation of HBV DNA and HBsAg levels in the four CHB phases. Each sample was represented by the blue dot according to their phase classification. Both the HBV DNA and HBsAg levels were calculated in $\log_{10}$ IU/mL. Correlation coefficient (r) and its significance (p) were shown in each phase. The correlation was considered strong for $r \geq 0.5$, moderate for $0.3 \leq r < 0.5$, and weak or negligible for $r < 0.3$. Significance was assessed at $p < 0.05$.

Table 1. Characteristics of chronic hepatitis B patients.

	Overall (n=152)	Immune-tolerance (n=33)	Immune-clearance (n=32)	Low/non replicative (n=34)	e- negative hepatitis B (n=53)	ANOVA *p* value
Sex (M:F)	102:50	16:17	22:10	20:14	44:9	0.006
Age (years)[a]	44 (14 - 80)	39 (21–80)	38 (14–74)	47 (25–69)	49 (26–68)	0.002
ALT (U/L)[a]	49 (13 - 1365)	36 (13–89)	126 (16–1320)	31.5 (19–97)	87 (17–1365)	<0.001
AST(U/L)[a]	41 (11 - 977)	29 (11–212)	129.5 (19–977)	27 (12–89)	73 (11–900)	<0.001
Platelet (10^3/mm^3)[a]	272 (35–488)	311 (154–432)	272.5 (56–488)	262.5 (56–441)	180 (35–472)	0.001
Albumin (g/dL)[a]	3.80 (1.8–4.4)	4.00 (3.3–4.4)	3.80 (1.9–4.3)	4.00 (3.0–4.3)	3.20 (1.8–4.3)	<0.001
INR[a]	1.00 (0.8–2.3)	0.95 (0.86–1.20)	0.99 (0.88–2.3)	0.92 (0.8–1.8)	1.10 (0.8–2.25)	<0.001
HBsAg (log_{10} IU/mL)[a]	3.25 (-1.30–4.99)	4.22 (-1.3–4.99)	3.34 (-1.3–4.99)	2.52 (-1.3–3.93)	3.37 (-1.3–4.65)	0.001
HBV DNA (log_{10} IU/mL)[a]	4.24 (0.04 - 8.14)	4.55 (0.78–7.82)	5.13 (0.44–8.04)	1.73 (0.04–3.87)	4.83 (0.71–8.14)	<0.001
HBsAg/HBV-DNA	0.74 (0.03–45.1)	0.86 (0.03–3.59)	0.60 (0.10–4.28)	1.1 (0.50–45.16)	0.64 (0.16–2.96)	0.029
HBeAg (PEIU/mL)[a]	451.99 (1.02–4500.38)	592.0 (1.02–2376.29)	236.65 (1.03–4500.38)	ND	ND	0.803
Genotype[b]						
B	78 (77.2%)	19 (76 %)	14 (58.3%)	20 (87 %)	25 (86.2%)	0.059
C	23 (22.8%)	6 (24 %)	10 (47.7%)	3 (13 %)	4 (13.8%)	
Subtype[b]						
Adw	82 (75.2%)	19 (76 %)	14 (56 %)	19 (79.2%)	30 (85.7%)	0.197
Adr	24 (22%)	6 (24 %)	10 (40 %)	4 (16.7%)	4 (11.4%)	
Ayw	3 (2.8%)	0	1 (4 %)	1 (4.2%)	1 (2.9%)	
BCP[b]						
Mutant	31 (29.0%)	4 (21.1 %)	4 (22.2 %)	7 (28.0 %)	16 (35.6%)	0.586
Wild type	76 (71.0%)	15 (78.9 %)	14 (77.8 %)	18 (72.0 %)	29 (64.4%)	

Table 1. continued

	Overall (n=152)	Immune-tolerance (n=33)	Immune-clearance (n=32)	Low/non replicative (n=34)	e- negative hepatitis B (n=53)	ANOVA *p* value
Precore[b]						
Mutant	32 (29.9%)	0	3 (16.7 %)	11 (44.0%)	18 (40 %)	0.003
Wild type	75 (70.1%)	19 (100%)	15 (83.3 %)	14 (56.0%)	27 (60 %)	
Cirrhosis[b]						
Cirrhosis	24 (15.8%)	0	4 (12.5%)	1 (2.9%)	19 (35.8 %)	<0.001
Non Cirrhosis	128 (84.2 %)	33(100%)	28 (87.5%)	33 (97.1%)	34 (64.2 %)	

[a]median (min–max) [b]number (percent of detected samples)

Table 2. Correlation between HBsAg, HBeAg, and HBV DNA levels as well clinical parameters.

Patient group	Parameter		*r**	*p*
• HBeAg-positive patients Overall (N=65)	HBsAg	HBeAg	0.481	0.005
		HBV DNA	0.661	<0.001
		Age	-0.385	0.002
		ALT	0.115	0.364
	HBeAg	HBV DNA	0.133	0.460
		Age	-0.218	0.224
		ALT	-0.068	0.706
	HBV DNA	Age	-0.300	0.015
		ALT	0.482	<0.001
• Immune Tolerance (N=33)	HBsAg	HBeAg	0.456	0.076
		HBV DNA	0.712	<0.001
		Age	-0.440	0.010
		ALT	-0.216	0.228
	HBeAg	HBV DNA	0.218	0.418
		Age	0.009	0.974
		ALT	0.016	0.952
	HBV DNA	Age	-0.379	0.029
		ALT	0.004	0.982
• Immune Clearance (N=32)	HBsAg	HBeAg	0.362	0.169
		HBV DNA	0.731	<0.001
		Age	-0.433	0.015
		ALT	0.635	<0.001
	HBeAg	HBV DNA	0.195	0.453
		Age	-0.477	0.053
		ALT	0.000	1.000
	HBV DNA	Age	-0.339	0.057
		ALT	0.613	<0.001
• HBeAg-negative patients Overall (N=87)	HBsAg	HBV DNA	0.586	<0.001
		Age	-0.085	0.443
		ALT	0.378	<0.001
	HBV DNA	Age	-0.090	0.407
		ALT	0.434	<0.001
• Low-replicative (N=34)	HBsAg	HBV DNA	0.288	0.098
		Age	-0.315	0.069
		ALT	0.135	0.446
	HBV DNA	Age	-0.259	0.140
		ALT	0.199	0.259
• HBeAg-negative hepatitis (ENH) (N=53)	HBsAg	HBV DNA	0.429	0.002
		Age	-0.020	0.892
		ALT	0.153	0.293
	HBV DNA	Age	-0.039	0.782
		ALT	0.100	0.482

*) Spearman Correlation Test

Overall, HBsAg had a significant negative correlation with age (r= -0.327; p<0.001). Based on HBeAg status, the correlation was found in the HBeAg-positive patients (r=-0.385; p=0.002), but not in the HBeAg-negative patients (r=-0.085; p=0.443). Accordingly, when analyzed separately for each CHB phase, this negative correlation was observed only in IT (r=-0.440; p=0.010) and IC (r=-0.433; p=0.015).

The HBsAg levels had weak correlations with AST (r=0.221; p=0.007) and ALT (r=0.264; p=0.001) in the 152 patients. Broken down by phase, it was demonstrated that there was a strong correlation between HBsAg levels and ALT in IC (r=0.635; p<0.001), but not in the other phases. We did not observe significant correlations between HBsAg and gender, nor with albumin, platelet and INR. Furthermore, the incidence of cirrhosis in 24 patients did not correlate with the HBsAg levels (r=0.106; p=0.202).

A significant negative correlation was also observed between HBV DNA levels and age in the entire patients (r=-0.233; p=0.004). However, analysis based on HBeAg status revealed that this correlation was found only in the HBeAg-positive patients (r=-0.300; p=0.015). When analyzed separately for each phase, this negative correlation was only noted in the IC phase (r=-0.379; p=0.029) (Table 2). According to clinical parameters, HBV DNA was positively correlated with AST (r=0.431; p<0.001), ALT (r=0.457; p<0.001), and INR (r=0.258; p=0.002), but negatively correlated with albumin (r=-0.231; p=0.005) and platelet (r=-0.170; p=0.039). Specifically by each phase, the correlation between HBV DNA and ALT was observed only in IC (r=0.613; p<0.001). In contrast to observation with HBsAg levels, the HBV DNA levels were correlated with the occurrence of cirrhosis (r=0.203; p=0.012).

HBV genotypes/subtypes and the presence of BCP A1762T/G1764A and precore G1896A mutations

A total of 101 patient samples were genotyped, resulting in 78 (77.2%) genotype B and 23 (22.8%) genotype C. HBeAg-negative patients had more genotype B compared with HBeAg-positive patients (p = 0.018). Subtype determination of 109 samples showed 82 (75.2 %) patients with subtype *adw*, 24 (22%) *adr* and 3 (2.8%) *ayw*. There was no significant difference in the distribution of genotype and subtype between CHB phases (p=0.059 and p=0.197, respectively) (Table 1). Genotyping and subtyping were not successfully performed in a subset of patients due to insufficient HBV DNA content. For similar reason, analysis for the presence of BCP and precore mutations was effectively done on sequences from 107 patients. The BCP and precore mutations were identified in 31 (29.0%) and 33 (29.9%) of the 107 patients, respectively (Table 1). There was no significant difference in the frequency of BCP mutations between HBeAg-positive and

HBeAg-negative groups. On the other hand, the precore mutations were significantly more prevalent in HBeAg-negative than in HBeAg-positive patients (41.4% vs 8.1%; p<0.001) with comparable occurrence in both genotypes B and C (p=0.151, data not shown). The precore mutations were found in all phases except IT, with the highest frequency in the LR (p=0.003) (Table 1).

An interesting BCP and precore mutation distribution pattern was observed when the prevalence was evaluated according to the patient age. The frequency of the BCP mutations increased with age in HBeAg-positive patients (r=0.345; p=0.037), but not in HBeAg-negative patients (r=-0.037; p=0.762). The precore mutation frequency did not show any correlation with patient age either in HBeAg-negative (r=-0.062; p=0.611) or in HBeAg-positive groups (r=-0.204; p=0.225). The distribution of the BCP and precore mutations according to the patient age-group and HBeAg status is presented in Fig. 2a.

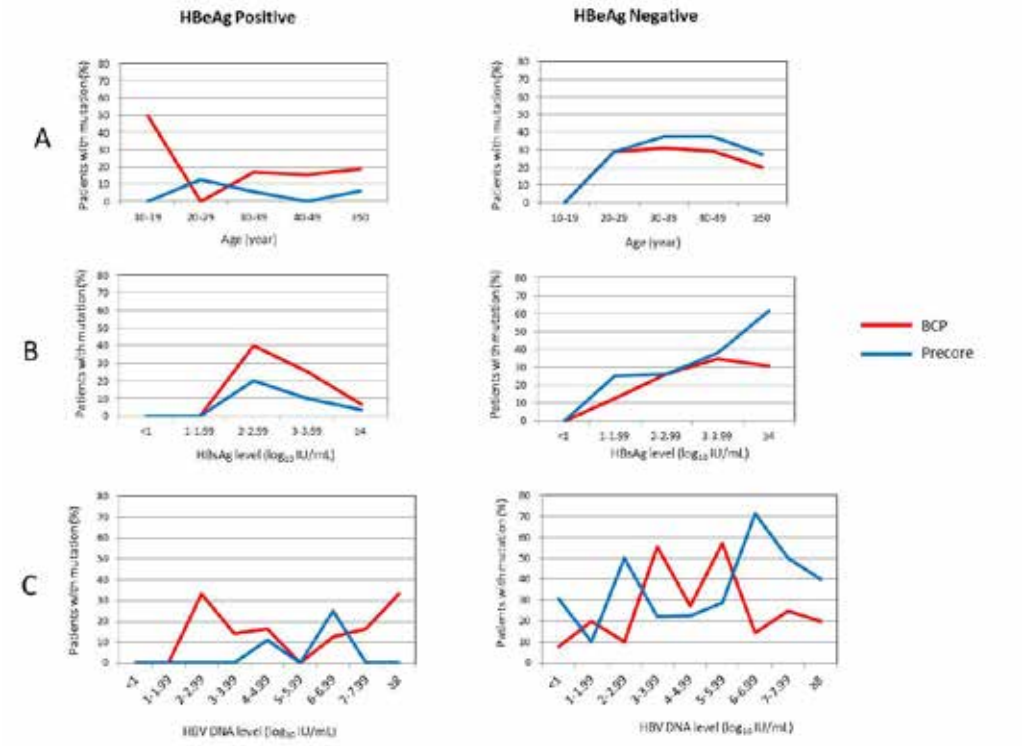


Figure 2. Prevalence of patients with mutations at BCP (red) or precore (blue) regions based on HBeAg status. The graphs showed prevalence of mutant samples according to (A) patient age groups, (B) HBsAg levels in $\log_{10}$ IU/mL, and (C) HBV DNA levels in $\log_{10}$ IU/mL.

We also analyzed the occurrence of the BCP and precore mutations based on HBeAg status in connection with HBsAg and HBV levels (Fig. 2b and 2c). The BCP mutations did not show any correlation with HBsAg and HBV DNA levels in both HBeAg-positive and HBeAg-negative groups, while the precore mutations was significantly correlated with HBsAg levels in the HBeAg-negative group only (r=0.341, p=0.005), with no correlation in the HBeAg-positive group (r=-0.185, p=0.272). Further analysis for each CHB phase revealed positive correlations between BCP mutations and HBsAg levels in LR (r=0.568, p=0.003) but negative correlation in IT (r=-0.660, p=0.002). The BCP mutations also showed negative correlation with HBV DNA in the IT phase (r=-0.660, p=0.002). A significant positive correlation was observed between precore mutations and HBsAg as well as HBV DNA levels in ENH (r=0.556; p<0.001 and r=0.313; p=0.036, respectively). In contrast, a negative correlation between precore mutations and HBV DNA levels was seen among IC patients (r=-0.477; p=0.046).

Table 3. Virological characteristics of the study subjects.

	Cirrhosis	Non cirrhosis	*p**
n	24	128	
Age #	50.2 (34–74)	41.8 (14–80)	*0.006*
HBsAg#	3.68 (-1.30–4.91)	3.19 (-1.30–4.99)	*0.202*
HBV DNA#	5.23 (0.71–8.04)	4.16 (0.04–8.14)	*0.012*
Genotype			
B	12 (85.7 %)	66 (75.9 %)	*0.333*
C	2 (14.3 %)	21 (24.1 %)	
Subtype			
adw	11 (78.6 %)	71 (74.7 %)	*0.453*
adr	2 (14.3%)	22 (23.2 %)	
ayw	1 (7.1%)	2 (2.1 %)	
BCP A1762T/G1764A			
Wild type	14 (63.6 %)	62 (72.9 %)	*0.272*
Mutant type	8 (36.4)	23 (27.1)	
Precore G1896A			
Wild type	14 (63.6 %)	61 (71.8 %)	*0.310*
Mutant type	8 (36.4 %)	24 (75 %)	

*) Mann-Whitney Test for continue data and Chi Square Test for categorical data
#) Data were express in median (min-max)

Subgroup analysis of patients with cirrhosis

This study included 24 (15.8 %) cirrhotic patients, distributed in the IC, LR and ENH phases. The median age of these patients was 50.2 years (range: 34-74 years), significantly older than the non-cirrhotic patients (median age 41.8 years; p=0.006). The median value of HBsAg levels was comparable between the cirrhotic and non-cirrhotic groups (3.68 log IU/mL vs 3.19 log IU/ml, p=0.202), while the median value of HBV DNA level was significantly higher in the cirrhotic group compared with that in the non-cirrhotic

group (5.23 $\log_{10}$ IU/mL vs 4.16 $\log_{10}$ IU/mL; Mann-Whitney test, p=0.012). A significant correlation between HBsAg and HBV DNA levels was observed in the cirrhotic group (r = 0.642, p<0.001).

There was no significant difference in the distribution of HBV genotypes (B and C) and subtypes (*adw*, *adr*, *ayw*) between the two patient groups in all CHB phases (χ^2 test; p=0.333 and p=0.453, respectively). The occurrence of BCP and precore mutations was also comparable in these two groups. Description of characteristics of cirrhotic and non-cirrhotic patients is shown in Table 3.

Discussion

Chronic hepatitis B is a dynamic process during which the course of the disease is determined by complex interactions between HBV, the infected hepatocyte and the host immune response.[8] In recent years, our understanding of the pathogenesis and natural history of CHB has been facilitated by the availability of assays quantifying HBsAg[25,26,27,28], and the recent development of a protocol for HBeAg quantitation[27], in addition to the established systems for measuring serum HBV DNA. This study aimed to investigate the HBsAg and HBeAg levels, HBV genotype, as well as the prevalence of BCP and precore mutations during the natural history of CHB patients. To our knowledge, this is the first study that evaluates quantitative HBV serology and HBV DNA characteristics in CHB patients in Indonesia.

This study showed that HBsAg levels differed significantly between the four phases of CHB. In agreement with previous reports[26,29,30,31,32], the HBsAg levels were significantly higher in HBeAg-positive compared to HBeAg-negative patients. Specifically in each phase, the median value of HBsAg levels was highest in the IT phase, decreased in the IC, became lowest in the LR, and increased again in the ENH. The declining order of HBsAg levels from IT, IC to LR can be explained that in the IT phase, which is the early phase of CHB, HBV virions and their antigens are not or minimally subjected to host immune reaction, characterized by normal ALT with no or minimal histologic activity.[33] As host immune system develops and matures, immune activation would cause the patients to enter the IC phase, characterized by continuing hepatitis activity or episodic flare with increase of ALT levels and histologic activity. These events may lead to a decrease in HBsAg levels, eventually followed by HBeAg clearance and appearance of its antibody (anti-HBe), defined as HBeAg seroconversion.[34]

HBeAg seroconversion is followed by sustained clinical remission with low or undetectable HBV DNA and normal ALT levels, marking the 'inactive carrier state' (LR) phase of CHB. The negative correlation between HBsAg and age documented in this study could be another evidence of effective immune control of HBV replication.[32] The increase of HBsAg levels accompanied by higher ALT in the ENH patients could suggest disease reactivation that triggers immune-mediated liver injury.[25,31,35] In the ENH, HBeAg-negative patients may develop flares and disease progression, comparable to HBeAg-positive patients in the IC phase.[26,34] Because the immunopathogenesis of ENH is similar to that of HBeAg-positive hepatitis, this reactive phase is also viewed as a variant of the IC, as adopted in the recent update of the APASL guidelines.[6,35]

HBV DNA levels also varied between the CHB phases. Significantly higher levels of HBV DNA were found in HBeAg-positive patients compared to HBeAg-negative patients. With a stratified analysis by phase, higher HBV DNA levels were evident in the IT, IC and ENH. The decrease of HBV DNA levels from IC to LR by a median value of more than 3 $\log_{10}$ could suggest successful viral immune clearance by the hosts.[8] The negative correlation between HBV DNA and patient age found in this study confirmed this finding. As similarly observed in the change of HBsAg, there was a rise of HBV DNA from LR to ENH, suggesting disease reactivation in HBeAg-negative patients.[31]

A good correlation between HBsAg and HBV DNA levels was revealed in the entire patients. Based on HBeAg status, there was a better correlation in HBeAg-positive patients as compared with HBeAg-negative patients. When analyzing the phases separately, the correlation was strong in the IT and IC, but missing in the LR, while a modest correlation was noted in the ENH. These results are encouragingly consistent with most studies, that the correlation between HBsAg and HBV DNA is mostly seen in the HBeAg-positive patients, particularly in the IT phase, and limited or absent in HBeAg-negative patients, namely in the LR phase.[30,31,32,36] This finding may reflect a relationship between HBsAg production and viral replication in the early phases of CHB, which appeared to be dissociated after the onset of immune clearance by the host immunity.[25]

The dissociation between HBsAg and HBV DNA levels was displayed by the ratio of HBsAg/HBV-DNA, which was stable in HBeAg-positive patients, but sharply increased after HBeAg seroconversion. A possible explanation for this finding is that these two viral markers are synthesized via different pathways and under different immune control mechanisms.[31,32] In addition to being the envelope of infectious HBV particles, HBsAg is also secreted as non-infectious filamentous or spherical particles.[29,31,37] Thus, HBsAg production far exceeds the amount required for virion assembly, with all three forms of the circulating HBsAg indifferently detected by HBsAg quantification assays. Another

possible explanation would be the production of HBsAg from segmented HBV sequences integrated into the host genome, which is a source for HBsAg production other than cccDNA.[30] All these indicate that immune control of HBV replication does not necessarily impair HBsAg production.

The reduction of HBsAg levels from IC to LR by a median value of less than 1 $\log_{10}$ IU/mL, which was much smaller than the decrement in HBV DNA levels, should be seen as important.[3] It may suggest that the production of HBsAg is well preserved, reflecting the preservation of cccDNA activity.[14] Studies have shown that patients are still at substantial risk of developing severe complications and HCC, even in the low level HBV viremia phase.[38,39] It has been reported that HBV reactivation occurs more frequently in inactive carriers in endemic regions than in Western countries.[40] This finding reiterates the importance of continuous monitoring in CHB patients.[41]

Despite the potential utility of HBsAg quantitation in clinical practice and its contribution in basic research, its application still needs to be carefully assessed. Recently, Hsu *et al*.[42] reported the emergence of S gene mutants in patients having HBsAg seroconversion after peginterferon therapy. Two patients remained viremic, with one of them experiencing an episode of hepatitis relapse seven months after the end of treatment and was negative for HBsAg.[42] More recently, Policino *et al*. found that Pre-S/S HBV variants, which are commonly found in CHB patients, may cause dissociation between HBsAg production and HBV DNA replication due to alterations in the pathway of the HBsAg synthesis resulting in the reduction of HBsAg level.[43] These reports may suggest the importance of studies on the clinical performance of quantitative HBsAg in different patient populations, especially in regions with high prevalence of CHB where diversity of HBV strains is common.

A correlation between HBsAg with both HBeAg and HBV DNA was observed in this study. However, such correlation was not seen between HBeAg and HBV DNA. This situation could be the consequence of the separate pathways for HBeAg production and viral replication. The discrepant correlation could also be due to the increase of viral variants defective for HBeAg production approaching seroconversion. The presence of such variants should be considered when developing treatment algorithms for monitoring disease progression, treatment efficacy and drug resistance that include quantitative HBeAg.[29]

Relating to clinical parameters, an overall correlation of HBsAg and HBV DNA with ALT was observed. Separately by phases, the correlation was found only in the IC, representing more active immunity against HBV. The fact that similar correlation was not seen in other phases could be contributed by the higher prevalence of cirrhotic patients

in HBeAg-negative group, especially in the ENH. The lower albumin and platelet levels as well as the higher INR in ENH patients supported this observation. In light of these findings, caution should be taken in assessing HBeAg-negative patients with normal or minimally increased ALT levels, particularly in populations in which acquisition of HBV occurs early in life. Several clinical studies indicate that significant liver damage can occur in asymptomatic CHB patients with high viral loads and persistently normal ALT levels.[44,45,46] These patients could be easily excluded from treatment as a result of ALT-based selection criteria for treatment initiation.[15] Another concern of the inaccuracy of ALT is the variability within the target population that is often not accounted for when categorizing patients by the ALT levels such as age, gender, body mass index (BMI), some metabolic parameters, as well as differences in the commercial assays used and the reference population chosen to establish the normal range.[15]

The genotype of HBV has been reported as a factor affecting HBsAg and HBV DNA levels in CHB[3,30,37] In this study, these viral markers were comparable among patients with genotype B and C. Positive correlations between HBsAg and HBV DNA were observed in both genotypes. We found that the prevalence of HBeAg was higher in patients with genotype C, suggesting that HBeAg seroconversion occurred at a lower rate among these patients. We also performed identification of HBsAg subtypes. The three commonly circulating subtypes in Indonesia[47,48] — *adw, adr, ayr* — were detected with comparable distribution in each phase.

Specific mutations in the HBV genome may occur during the course of chronic infection, with the BCP and precore mutations considered as the most important mutations associated with progression of liver disease.[49,50] The BCP mutations have been identified to decrease the binding of transcriptional factors causing relative reduction in HBeAg production[51], while the precore mutation encodes for a stop codon that abolishes the production of HBeAg.[29,52] Several studies showed that HBV variants could be detected with low percentage in the early phase of CHB and remain stable for years, but sharply increase in the years prior to HBeAg seroconversion.[29,49] After seroconversion, the variant percentage could remain high or increase in the majority of patients.[17,49]

The progressive increase in viral diversity, with selection of quasispecies defective for HBeAg production, has been assumed to be driven by the host immune pressure during the viral evolution.[29] The pattern of this dynamics was clearly demonstrated in the current study, with findings that the BCP mutations were correlated with ALT and age, particularly in the HBeAg-positive patients, as well as the progressively increasing prevalence of the precore mutations that reached its highest levels after seroconversion. The negative correlations between the BCP mutations with HBsAg and HBV DNA in the

IT, as well as the precore mutations with HBV DNA in the IC, may show the domination of the wild type virus in the early phase of CHB.[49] This finding is supported by other report from Wang *et al.,* that during early phase of CHB, the lack of host immune activity drives the increase of viral copy numbers, especially those with wild type genomes.[52] Conversely, positive correlations between the BCP mutations with HBsAg in the LR and the precore mutations with HBV DNA in the ENH may represent the 'replication advantage' of these variants over the wild type viruses in the later stages of CHB where host immune pressure has taken place.[49,52,53] This finding could be a sign that there is a temporal association between HBeAg seroconversion and the increasing prevalence of these mutants, which have prevailed in the early phase of CHB and accumulated as the disease progresses.

BCP mutations have been associated with the progression of liver disease and higher incidence of HCC. In this study, the BCP mutations were found in comparable distribution among the phases of CHB. This could suggest that the deleterious effect of these mutants may occur at any time during the course of the disease, implying the importance of continuous monitoring of CHB patients.[50,54,55] The precore mutations were found in escalating prevalence by CHB phases with higher increments than the BCP mutations. This may be due to the faster rate of selection that applies for precore mutations compared to the BCP mutations, because the precore mutants do not produce HBeAg at all and may have better survival advantage than the BCP mutants.[49]

At present, the role of precore mutations in the progression of liver disease remains unclear. Reports from Europe indicated that precore mutations were associated with progressive HBeAg-negative chronic hepatitis.[56,57] However, reports from Asia indicated that the acquisition of these mutants was related to inactive liver disease and detected at similar percentages in asymptomatic carriers as well as in patients with cirrhosis and HCC[19,58,59] Here, we observed that the occurrence of the precore and BCP mutants was comparable in cirrhotic and non-cirrhotic patients. Further studies with larger number of patients are needed to elucidate the role of precore mutations in CHB patients, stratified by different ethnic and patient populations as well as HBV genotypes.

We acknowledge several limitations in this study. Firstly, the cross-sectional design of this study might not reflect the highly dynamic nature of CHB, because the available data were based on single measurements that could not represent the fluctuating profile and the viral evolution during the natural history of HBV infection. As for treatment-naïve patients, particularly in the IC or ENH phases, follow-up studies without treatment for a long period would not be carried out, because they are the best candidates for active antiviral treatment.[6,30,60] Longitudinal pre- and on-treatment studies in cohorts of patients

with different CHB phases are necessary, stratified according to the host background and the viral genetic characteristics. Secondly, the fluctuating ALT levels of CHB may misclassify the phase categories in the natural history of this disease, because we only measured the ALT in single times in this cross-sectional study. Thirdly, liver biopsies were not obtained in this study. It has been reported that the use of ALT and HBV DNA levels without resorting to liver biopsy may miss histologically significant disease in approximately 10% of the patients.[40] As the absolute levels of HBsAg and HBV DNA in serum may not always accurately reflect the state of the disease in the liver, analysis and quantification of intrahepatic HBV replicative forms—in particular cccDNA and pregenomic RNA (pgRNA)—need to be conducted to gain further insights into the natural history of hepatitis B as well as additional information for evaluating antiviral therapy.

In conclusion, this study demonstrates that: (i) The levels of HBsAg and HBV DNA differ significantly across the phases of CHB. The HBsAg levels were highest in the IT and lowest in the LR, while the HBV DNA levels were higher in the IT and IC phases and lowest in the LR. Increase of both markers was also observed in the ENH. (ii) The correlation between HBsAg and HBV DNA was strong in the IT and IC, modest in the ENH, but missing in the LR. The dissociation between HBsAg and HBV DNA synthesis after HBeAg seroconversion provides evidence that immune control of HBV replication does not necessarily impair HBsAg production. (iii) A correlation between HBsAg and HBV DNA was documented in the ENH, suggesting that the concept of HBsAg and HBV DNA disconnection in later phases of HBV infection is not automatically applicable in HBeAg-negative hepatitis. This finding supports the view that the ENH is a variant of the IC. (iv) BCP mutations were found in comparable distribution among CHB phases, implying that the deleterious impact of these mutations could occur at any time in the disease course. Meanwhile, the precore mutations were identified with increased prevalence as the phases progressed, and significantly highest in the LR. We also found a temporal association between HBeAg seroconversion and the increase of the prevalence of patients with variants defective for HBeAg production, suggesting the role of host immunity in driving the evolution of these variants. (v) Caution still needs to be taken in utilizing quantitative HBeAg in clinical setting, considering the presence of variants that may reduce or abolish HBeAg production. Overall, this study suggests that despite the wide use of quantitative HBsAg and the attractiveness of quantitative HBeAg, further evaluation of their performance in clinical practice are needed before including them in the treatment algorithm. Pre- and on-treatment studies are necessary, and these studies should be stratified based on the host ethnic background and the viral genetic characteristics, including genotypes and mutations.

Acknowledgements

This study was partly supported by a research grant from the Ministry for Research and Technology, Republic of Indonesia. The authors would like to thank Dr. Esther Ramono, MM, Director of Mitra Kemayoran Hospital, Dr. Djoko Kurdo and Dr. Wijanto from PT Frisian Flag Indonesia, and Dr. Christian Dwi Aryanto from PT Toyota Astra Indonesia for sample referral; Prodia Laboratory for HBsAg quantification and PT Roche Diagnostics Indonesia for HBeAg quantification. We sincerely thank all the patients for participation in this study.

References

1. World Health Organization. Hepatitis B. Fact sheet No. 204. Revised July 2012. Available at:http://www.who.int/mediacentre/factsheets/fs204/en. Accessed on August 1, 2012.
2. Khan M, Dong JJ, Acharya SK, Dhagwahdorj Y, Abbas Z, Jafri SMW, Mulyono DH, Tozun N, Sarin SK. Hepatology issues in Asia: Perspectives from regional leaders. J Gastroenterol Hepatol 2004;19:S419-S430.
3. Nguyen T, Desmond P, Locarnini S. The role of quantitative hepatitis B serology in the natural history and management of chronic hepatitis B. Hepatol Int 2009;3:S5–S15.
4. Laras A, Koskinas J, Dimou E, Kostamena A, Hadziyannis SJ. Intrahepatic level and replicative activity of covalently closed circular hepatitis B virus DNA in chronically infected patients. Hepatology 2006;44:694–702.
5. Beck J, Nassal M. Hepatitis B virus replication. World J Gastroenterol 2007;13(1):48-64.
6. Liaw YF, Kao JH, Piratvisuth T, Chan, HLY, Chien RN, Liu JC et al. Asian-Pacific consensus statement on the management of chronic hepatitis B: a 2012 update. Hepatology Int 2012. DOI 10.1007/s12072-012-9365-4.
7. Pungpapong S, Kim WR, Poterucha JJ. Natural history of hepatitis B virus infection: An update for clinicians. Mayo Clin Proc 2007;28(8):967–975.
8. Liaw YF. HBeAg seroconversion as an important end point in the treatment of chronic hepatitis B. Hepatol Int 2009;3:425–433.
9. European Association for the Study of the Liver. EASL clinical practice guidelines: management of chronic hepatitis B virus infection. J Hepatology 2012; 57:167–185.
10. Fang ZL, Sabin CA, Dong BQ, Wei SC, Chen QY, Fang KX, et al. The association of HBV core promoter double mutations A1762T and G1764A with viral load differs between HBeAg positive and anti-HBe positive individuals: a longitudinal analysis. J Hepatol 2009;50:273–280.

11. Yim HJ, Lok AS. Natural history of chronic hepatitis B virus infection: what we knew in 1981 and what we know in 2005. Hepatology 2006;43(2 Suppl 1):S173-81.

12. Liaw YF, Leung N, Kao JH, Piratvisuth T, Gane E, Han KH et al. Asian–Pacific consensus statement on the management of chronic hepatitis B: a 2008 update. Hepatol Int 2008;2:263–283.

13. Tong S, Kim KH, Chante C, Wands J, Li J. Hepatitis B virus e antigen variants. Int J Med Sci 2005;2(1):2–7.

14. Werle B, Cinquin K, Marcellin P, Pol S, Maynard M, Trépo C, Zoulim F. Evolution of hepatitis B viral load and viral genome sequence during adefovir dipivoxil therapy. J Viral Hepat 2004;11(1):74–83.

15. Ahn SH, Chan HLY, Chen PJ, Cheng J, Goenka MK, Hou J, et al. Chronic hepatitis B: whom to treat and for how long? Propositions, challenges, and future directions. Hepatol Int 2010;4:386–395.

16. Sonneveld MJ, Zoutendijk R, Janssen HLA. Hepatitis B surface antigen monitoring and management of chronic hepatitis B. J Vir Hepat 2011;18:449–457.

17. Chu CJ, Hussain M, Lok ASF. Quantitative serum HBV DNA levels during different stages of chronic hepatitis B infection. Hepatology 2002;36:1408–1415.

18. Shi YH, Shi CH. Molecular characteristics and stages of chronic hepatitis B virus infection. World J Gastroenterol 2009;15(25):3099–3105.

19. Tong MJ, Hsu L, Hsien C, Kao JH, Durazo FA, Saab S, Blatt LM. A comparison of hepatitis B viral markers of patients in different clinical stages of chronic infection. Hepatol Int 2010;4:516–522.

20. Okamoto H, Nishizawa T. Non-B Non-C Non-G hepatitis virus gene, polynucleotide, polypeptide, virion, method for separating virion, and method for detecting virus; 1992. http://www.freepatentsonline.com/EP1010759.html (with permission).

21. Iizuka H, Ohmura K, Ishijima A, Satoh K, Tanaka T, Tsuda F, Okamoto H, Miyakawa Y, Mayumi M. Correlation between anti-HBc titers and HBV DNA in blood units without detectable HBsAg. Vox Sang 1992;63:107–111.

22. Quarleri J, Moretti F, Bouzas MB, Laufer N, Carrillo MG, Giuliano SF, Pérez H, Cahn P, Salomon H. Hepatitis B virus genotype distribution and its lamivudine-resistant mutants in HIV-coinfected patients with chronic and occult hepatitis B. AIDS Res Hum Retroviruses 2007;23(4):525–531.

23. Okamoto H, Imai M, Tsuda F, Tanaka T, Miyazawa Y, Mayumi M. Point mutation in the S gene of hepatitis B virus for a d/y or w/r subtypic change in two blood donors carrying a surface antigen of compound subtype adyr or adwr. J Virol 1987;61(10):3030–3034.

24. Norder H, Courouce AM, Magnius LO. Molecular basis of hepatitis B virus serotype variations within the four major subtypes. J Gen Virol 1992;73:3141–3145.

25. Chan HLY, Wong VWS, Wong GLH, Tse CH, Chan HY, Sung JJY. A longitudinal study on the natural history of serum hepatitis B surface antigen changes in chronic hepatitis B. Hepatology 2010;52:1232–1241.

26. Liaw YF. Clinical utility of hepatitis B surface antigen quantitation in patients with chronic hepatitis B: a review. Hepatology 2011;53:2121–2129.

27. Wursthorn K, Zacher BJ, Jaroszewicz J, Darnedde M, Manns M, Wedemeyer H. Development of a protocol for the quantitative determination of HBeAg using the Elecsys® HBeAg immunoassay. J Viral Hepat 2011;18(7):179–183.

28. Tuaillon E, Mondain AM, Nagot N, Ottomani L, Kania D, Nogue E, Rubbo PA, Pageaux GP, Van de Perre P, Ducos J. Comparison of serum HBsAg quantitation by four immunoassays, and relationships of HBsAg level with HBV replication and HBV genotypes. PLoS One 2012;7(3):e32143.

29. Thompson AJ, Nguyen T, Iser D, Ayres A, Jackson K, Littlejohn M, Slavin J, Bowden S, Gane EJ, Abbott W, Lau GK, Lewin SR, Visvanathan K, Desmond PV, Locarnini SA. Serum hepatitis B surface antigen and hepatitis B e antigen titers: disease phase influences correlation with viral load and intrahepatic hepatitis B virus markers. Hepatology 2010;51(6):1933–1944.

30. Jaroszewicz J, Serrano BC, Wursthorn K, Deterding, Schlue J, Raupach R, Flisiak R, Bock CT, Manns MP, Wedemeyer H, Cornberg M. Hepatitis B surface antigen (HBsAg) levels in the natural history of hepatitis B virus (HBV)-infection: a European perspective. J Hepatol 2010;52:514–522.

31. Nguyen T, Thompson AJV, Bowden S, Croagh C, Bell S, Desmond PV, Levy M, Locarnini SA. Hepatitis B surface antigen levels during the natural history of chronic hepatitis B: a perspective on Asia. J Hepatol 2010;52:508–513.

32. Kim YJ, Cho HC, Choi MS, Lee JH, Koh KC, Yoo BC, Paik SW. The change of the quantitative HBsAg level during the natural course of chronic hepatitis B. Liver Int 2011;6:817–823.

33. Hwang S, Ahn CS, Song GW, Kim KH, Moon DB, Oh HB, Lim YS, Lee HC, Ha TY, Jung DH, Chung YH, Lee SG. Posttransplantation prophylaxis with primary high-dose hepatitis B immunoglobulin monotherapy and complementary preemptive antiviral add-on. Liver Transpl 2011;17(4):456–465.

34. Liaw YF. Hepatitis flares and hepatitis B e antigen seroconversion: implication in anti-hepatitis B virus therapy. J Gastroenterol Hepatol 2003;18(3):246–252.

35. Chu CM, Liaw YF. Hepatitis B surface antigen seroclearance during chronic HBV infection. Antivir Ther 2010;15(2):133–143.

36. Moucari R, Korevaar A, Lada O, Martinot-Peignoux M, Boyer N, Mackiewicz V, Dauvergne A, Cardoso AC, Asselah T, Nicolas-Chanoine MH, Vidaud M, Valla D, Bedossa P, Marcellin P. High rates of HBsAg seroconversion in HBeAg-positive chronic hepatitis B patients responding to interferon: a long-term follow-up study.J Hepatol 2009;50(6):1084–1092.

37. Brunetto MR. A new role for an old marker, HBsAg. J Hepatol 2010;52:475–477.

38. Mendy ME, Welzel T, Lesi OA, Hainaut P, Hall AJ, Kuniholm MH et al. Hepatitis B viral load and risk for liver cirrhosis and hepatocellular carcinoma in The Gambia, West Africa. J Viral Hepat 2010;17:115–122.

39. Chen JD, Yang HI, Iloeje UH, You SL, Lu SN, Wang LY, Su J, Sun CA, Liaw YF, Chen CJ. Carriers of inactive hepatitis B virus are still at risk for hepatocellular carcinoma and liver-related death. Gastroenterology 2010;138(5):1747–1754.

40. Villa E, Fattovich G, Mauro A, Pasino M. Natural history of chronic HBV infection: special emphasis on the prognostic implications of the inactive carrier state versus chronic hepatitis. Dig Liver Dis 2011;43 (Suppl 1):S8–14.

41. Ganem D, Prince AM. Hepatitis B virus infection-natural history and clnical consequences. N Engl J Med 2004;350(11):1118–1129.

42. Hsu CW, Yeh CT. Emergence of hepatitis B virus S gene mutants in patients experiencing hepatitis B surface antigen seroconversion after peginterferon therapy. Hepatology 2011;54(1):101–108.

43. Pollicino T, Amaddeo G, Restuccia A, Raffa G, Alibrandi A, Cutroneo G, Favaloro A, Maimone S, Squadrito G, Raimondo G. Impact of hepatitis B virus (HBV) preS/S genomic variability on HBV surface antigen and HBV DNA serum levels. Hepatology 2012;56(2):434–443.

44. Kumar M, Chauhan R, Gupta N, Hissar S, Sakhuja P, Sarin SK. Spontaneous increases in alanine aminotransferase levels in asymptomatic chronic hepatitis B virus-infected patients. Gastroenterology 2009;136(4):1272–1280.

45. Tsang PS, Trinh H, Garcia RT, Phan JT, Ha NB, Nguyen H, Nguyen K, Keeffe EB, Nguyen MH. Significant prevalence of histologic disease in patients with chronic hepatitis B and mildly elevated serum alanine aminotransferase levels. Clin Gastroenterol Hepatol 2008;6(5):569–574.

46. Park JY, Park YN, Kim DY, Paik YH, Lee KS, Moon BS, Han KH, Chon CY, Ahn SH. High prevalence of significant histology in asymptomatic chronic hepatitis B patients with genotype C and high serum HBV DNA levels. J Viral Hepat 2008;15(8):615–621.

47. Mulyanto, Depamede SN, Surayah K, Tsuda F, Ichiyama K, Takahashi M, Okamoto H. A nationwide molecular epidemiological study on hepatitis B virus in Indonesia: identification of two novel subgenotypes, B8 and C7. Arch Virol 2009;154(7):1047–1059.

48. Nurainy N, Muljono DH, Sudoyo H, Marzuki S. Genetic study of hepatitis B virus in Indonesia reveals a new subgenotype of genotype B in east Nusa Tenggara. Arch Virol 2008;153(6):1057–1065.

49. Nie JJ, Sun KX, Li J, Wang J, Jin H, Wang L, Lu FM, Li T, Yan L, Yang JX, Sun MS, Zhuang H. A type-specific nested PCR assay established and applied for investigation of HBV genotype and subgenotype in Chinese patients with chronic HBV infection. Virol J 2012;9(1):121.

50. Chu CJ, Keeffe EB, Han SH, Perrillo RP, Min AD, Soldevila-Pico C, Carey W, Brown RS Jr, Luketic VA, Terrault N, Lok AS. Prevalence of HBV precore/core promoter variants in the United States. Hepatology 2003;38(3):619–628.

51. Laras A, Koskinas J, Hadziyannis SJ. In vivo suppression of precore mRNA synthesis is associated with mutations in the hepatitis B virus core promoter. Virology 2002;295(1):86–96.

52. Wang HY, Chien MH, Huang HP, Chang HC, Wu CC, Chen PJ, Chang MH, Chen DS. Distinct hepatitis B virus dynamics in the immunotolerant and early immunoclearance phases. J Virol 2010;84(7):3454–3463.

53. Carman WF, Jacyna MR, Hadziyannis S, Karayiannis P, McGarvey MJ, Makris A, Thomas HC. Mutation preventing formation of hepatitis B e antigen in patients with chronic hepatitis B infection. Lancet 1989;2(8663):588–591.

54. Kao JH, Chen PJ, Lai MY, Chen DS. Basal core promoter mutations of hepatitis B virus increase the risk of hepatocellular carcinoma in hepatitis B carriers. Gastroenterology 2003;124(2):327–334.

55. Tong MJ, Blatt LM, Kao JH, Cheng JT, Corey WG. Precore/basal core promoter mutants and hepatitis B viral DNA levels as predictors for liver deaths and hepatocellular carcinoma. World J Gastroenterol 2006;12(41):6620–6626.

56. Hunt CM, McGill JM, Allen MI, Condreay LD. Clinical relevance of hepatitis B viral mutations. Hepatology 2000;31(5):1037–1044.

57. Tong S, Kim KH, Chante C, Wands J, Li J. Hepatitis B virus e antigen variants. Int J Med Sci 2005;2(1):2–7.

58. Yuen MF, Lai CL. Natural history of chronic hepatitis B virus infection. J Gastroenterol Hepatol 2000;15 (Suppl):E20–24.

59. Shi YH, Shi CH. Molecular characteristics and stages of chronic hepatitis B virus infection. World J Gastroenterol 2009;15(25):3099–3105.

60. Lok AF, McMahon BJ. Chronic hepatitis B: update 2009. Hepatology 2009;50(3):661–662.

CHAPTER 7

Hepatitis B virus genotype determination using simple restriction fragment length polymorphism (RFLP) for the detection of diagnostic single nucleotide polymorphism (SNP) in the preS/S region

Meta D. Thedja[1,2], Greet Boland[2], Susan I. Ie[1], Turyadi[1], Jan Verhoef[2], and David H. Muljono[1,3,4]

[1]Eijkman Institute for Molecular Biology, Jakarta, Indonesia

[2]Eijkman Winkler Institute, Utrecht Medical Centre, Utrecht, Netherlands

[3]Faculty of Medicine, Hasanuddin University, Makassar, Indonesia

[4]Sydney Medical School, University of Sydney, Australia

1. Introduction

Hepatitis B Virus (HBV) has high genetic variability within its 3.2 kb genome, resulting into eight genotypes, genotype A through H. This classification is based on 8% or more divergence of the complete genome sequence, with each genotype has a distinct geographical and ethnic distribution worldwide. In addition, some subgenotypes have been distinguishable based on more than 4% genome diversity within the same genotype strains (Norder et al., 2004; Kramvis et al., 2005).

There is growing evidence in support of the role of HBV genotypes in the activity and progression of the disease and response to treatment. Several studies from Asia, where genotype B and C are predominant (Nurainy et al., 2008; Mulyanto et al., 2010; Thedja et al., 2011), have shown the correlation between HBV genotype and liver disease (Kao et al., 2000; Ding et al., 2001; Orito et al., 2001; Zeng et al., 2005; Nguyen et al., 2009). Genotype C has been found to be associated with higher incidence of cirrhosis and hepatocellular carcinoma, with less responsiveness to interferon compared to genotype B (Kao et al., 2000; Ding et al., 2001; Yuen et al., 2003; Liu et al., 2006). Long-term follow-up studies have shown that the cumulative rate of spontaneous HBeAg seroconversion was significantly higher in patients infected with genotype B compared to those infected with genotype C (Chu et al., 2002; Sumi et al., 2003; Chan et al., 2003). Studies from Europe found that genotype A was associated with higher rate of chronic HBV infection and more severe liver diseases compared to genotype D (Mayerat et al., 1999). However, sustained response to interferon and HBsAg clearance in patients infected with genotype A was higher compared to those with genotype D (Sanchez-Tapias et al., 2002; Erhardt et al., 2006).

Assessing the practical applications as described above, it is likely that HBV genotyping should preferably be performed prior treatment, to estimate the success-rate of therapy and prognosis of the disease. Whole genome sequencing as the golden standard of HBV genotyping is not always feasible due to the cost and time constraints in analyzing large number of samples, particularly in developing countries. Several HBV genotyping methods have been reported, including PCR-based assays (Naito et al., 2001), enzyme-linked immunoassay (Usuda et al, 1999; Usuda et al., 2000), restriction fragment length polymorphism (Lindh et al, 1998; Mizokami et al, 1999; Aslam et al, 2008), line probe assay (Osiowy and Giles, 2003), real time PCR (Liu et al, 2006), and HBV DNA-chip assay (Pas et al, 2008).

Previous studies reported Single Nucleotide Polymorphisms (SNPs) within the HBV preS2 sequence that are useful for genotyping and subgenotyping purposes, in particular

for the Asian specific genotypes B, C, and D (Nurainy et al., 2008; Thedja et al., 2011). We carried out this study to establish a simple and efficient RFLP PCR-based assay for HBV genotyping involving the highly variable sequence of the complete preS2 region. The method was also confirmed by two other methods: sequencing and analysis of SNPs diagnostic site of the same region. The RFLP method successfully genotyped HBV strains into eight genotypes, providing a useful tool for rapid and inexpensive genotyping, particularly for endemic regions in Asia.

Table 1. HBV genotypes/subgenotypes and their restriction site digested by the two principal restriction enzymes, BamH1 and HPyCH4V. Some isolates showed the same digestion pattern that they should be followed with AluI or BmrI.

Genotype/ subgenotype	No. of sequence	Restriction site and restriction pattern digested with (bp)				Note
		BamHI (nt)	HpyCH4V (nt)	AluI (nt)	BmrI (nt)	
A1	15	(-) 302	(3208, 26) 85 33 184	(8) 101 201		Uncut group
B	54	(-) 302	(3208) 85 217			
C	28	(-) 302	(3208, 147) 85 154 63		(137) 229 73	
D1/D4	16	(-) 302	(3208, 26) 85 33 184	(-) 302		
D2/D3	7	(-) 302	(3132, 3208, 26, 147) 8 77 33 121 63			
G	5	(-) 302	(3208, 147) 85 154 63		(-) 302	
A2	9	(29) 121 181	(3208, 26) 85 33 184			Cut group
E	13	(29) 121 181	(3208, 26, 147) 85 33 121 63			
F	7	(29) 121 181	(3208) 85 217		(21) 113 189	
H	7	(29) 121 181	(3208) 85 217		(-) 302	

2. Materials and methods

2.1. Study on phylogenetic analysis and prediction of enzyme restriction sites.

HBV sequences were retrieved from the public database GenBank then aligned for construction of phylogenetic tree using BioEdit 5.0.9 and MEGA4 software. One hundred and twenty-seven of HBV complete genome sequences with known genotypes were used in this study, comprising 19 sequences of genotype A, 34 genotype B, 25 genotype C, 20 genotype D, 13 genotype E, 4 genotype F, 5 genotype G and 7 genotype H; and

AY2266578 (WMHBV) as outgroup. A phylogenetic tree was constructed with neighbor joining method and Kimura2-parameter to calculate the genetic distance, with 1000 bootstraps to get a more reliable genotype clustering. In parallel, the 302 bp (nt 3125 – nt 210) preS/S sequences obtained from the same sequences were aligned and constructed as a phylogenetic tree as well. Each of these 127 HBV sequences was further analyzed for identification of enzyme restriction site prediction for the eight HBV genotypes using the preS/S sequence (Table 1).

2.2. Samples.

One hundred serum samples from patients infected with hepatitis B virus, originated from various Asian and European ethnic origins, were collected in two countries: the Netherlands (79 samples) and Suriname (21 samples). Informed consent was obtained from each patient, and all the procedures were approved by the local Ethics Committee. The sample preparations and laboratory research were done in Eijkman Winkler Institute UMC, Utrecht, Netherlands.

2.3. Amplification, digestion reaction, sequencing, and confirmation by SNPs of preS/S region.

DNA was extracted from 140-µL serum using Viral DNA QIAquick spin column method extraction kit (Qiagen Inc., Chatsworth, CA) according to the manufacturer's instruction. Fragment preS/S was amplified using specific primer sets in a hemi-nested PCR (Takahashi et al., 1998). Primers used were PS1-1 (CCTCCTGCCTCCACCAATCG)/T703 (CAGAGTCTAGACTCGTGGTT) and PS5-2 (CTCGTGTTACAGGCGGGGTT)/T703, with detection of amplification products by 2% agarose gel electrophoreses. Restriction digestions of the RFLP were carried out as follows: 5µL of the purified PCR products, 1µL 1x Buffer B, 3.7µL ddH2O, and addition of 0.3µL BamHI or HPyCH4V in separate reactions, were mixed and incubated at 37°C for 5 hours. The digested PCR products were electrophoresed on 3% Nusieve agarose gel in ½ x TBE buffers containing ethidium bromide (10 mg/ml), and visualized under UV light. The genotype/subgenotype was determined by their distinct length sizes in gel electrophoresis. The reactions can be followed by other restriction enzymes, AluI or BmrI, depending on the pattern obtained. Direct sequencing reaction of the 100 purified preS/S PCR products was done to evaluate and confirm the validity of the RFLP genotyping method. Evaluation and confirmation were also done by employing specific Single Nucleotide Polymorphism (SNPs) analysis within the preS/S region as reported

previously (Nurainy et al., 2008; Thedja et al., 2011), particularly for genotype B, C, and D.

2.4. HBV DNA cloning for samples with unidentifiable genotype/subgenotype pattern.

Cloning was done for samples with non-specific or unidentified pattern. The purified second PCR products of the preS/S region (302 bp) were ligated into pGEMT-vector (Promega, Madison, USA), incubated, and cloned into DH5α competent cells. Ten positive clones from each sample were selected and amplified by PCR method using internal primers PS5-2 and T703, and sequenced as described previously.

3. Results

3.1. Phylogenetic analysis and characteristic of restriction site prediction of eight HBV genotypes.

Phylogenetic tree constructed based on the 127 HBV complete genomes was shown in Figure 1a. The tree showed the identification of the eight genotypes, A through H, with good bootstrap value for each of the major cluster. Phylogenetic tree constructed based on the preS/S sequences from the same samples showed complete agreement in the isolate clustering with that of the complete genome (Figure 1b), with the exception of one isolate of subgenotype D4 that grouped into subgenotype D1.

The restriction enzyme sites and specific restriction pattern of each HBV genotype were shown in Figure 2, with algorithm strategy for HBV genotyping in Figure 3. Four restriction enzymes: BamHI (G'GATC), HPyCH4V (TG'CA), AluI (AG'CT), and Bmrl (ACTGGG) were identified, and can be used in a three-step-method for genotyping/subgenotyping purposes. Firstly, BamHI was used to separate all the sequences into two groups: the cleaved group consisting of genotype A2, E, F, and H with the restriction site GGATC sequence at nt 3225-3226; and the uncut group of the other five genotypes (A1, B, C, D, and G) that did not have the restriction site within the preS/S sequence.

In the second step, each genotype of these two initial groups showed different patterns when subjected to digestion reactions with HpyCH4V enzyme due to the different positions of the restriction site TGCA sequence. Of the cleaved group, genotype A2 has two restriction sites at nt 3209-3210 and nt 3242-3243; E has three sites at 3209-3210,

nt 47-48, and nt 147-148; while the two genotypes, F and H, showed the same pattern with one restriction site at nt 3209-3210. These two genotypes were further treated with AluI enzyme in the third step. The presence of one AluI restriction site in genotype F produced two bands, distinguishing this genotype from genotype H that did not have the AluI site. Similarly, genotype A1 and D (D1 and D4) of the uncut group had HpyCH4V restriction sites at nt 3209-3210 and nt 47-48 at the same positions. Digestion with AluI enzyme distinguished genotype A1 from the genotype D (D1/D4) that did not have AluI site. In the third step for uncut group, BmrI enzyme was used to differentiate genotype C from G. Genotype C had BmrI restriction site while genotype G did not.

3.2. Determination of HBV genotype/subgenotype by specific RFLP pattern, and verification by sequencing and SNPs.

Analysis of RFLP pattern of the 100 isolates from the Netherlands and Suriname successfully determined HBV genotypes/subgenotypes in 97 (97%) isolates, while three isolates showed unclassifiable RFLP pattern (Figure 4). Seven HBV genotypes/subgenotypes: A, B, C, D, E, G and H were identified. Genotype A was found in 22 (22.7%) isolates (A1 in 7 and A2 in 15); B in 28 (28.9%); C in 19 (19.6%); D in 23 (23.7%); E in 3 (3.1%); G in 1; and H in 1 isolate. A phylogenetic tree was constructed by preS/S sequence of these isolates for genotype clustering (Figure 5). All one hundred isolates were clustered into seven HBV genotypes: A found in 23 (23%) (A1 in 5 and A2 in 18); B in 27 (27%); C in 23 (23%); D in 22 (22%); E in 3; G in 1; and H in 1 isolate.

Excluding the three isolates unclassifiable by RFLP, the data showed that genotype classification by RFLP and phylogenetic tree analysis agreed in 91 (93.8%) isolates as shown in Figure 5, while the other six isolates showed inconsistent results. Two isolates (7544 and 4435) identified as genotype A1 by RFLP clustered into genotype A2 by phylogenetic tree analysis. One isolate (Sur27) grouped as genotype B by RFLP, but clustered into genotype C by phylogenetic tree analysis. Two isolates were identified as genotype D by RFLP but grouped into genotype A1 (38826) and C (998) by phylogenetic tree analysis. Another isolate (28159) identified as genotype B by RFLP, and genotype C by phylogenetic tree. These six isolates and one isolate (6809) unclassifiable by RFLP were further analyzed by cloning and sequencing, while the other two isolates unclassifiable by RFLP (Sur4 and 5002) were excluded due to deletion of 43 and 49 nucleotides as found in initial sequencing, respectively.

Since previous reports of SNPs diagnostic sites characterized only three genotypes: HBV/B, C, and D (Nurainy et al., 2008; Thedja et al., 2011), the three isolates unclassifiable by RFLP and 27 isolates with non B/C/D genotype determined by RFLP

and phylogenetic tree analysis were excluded. Of a total 70 isolates with HBV/B, C, and D genotypes, 56 (80%) had identical genotyping results with the three methods: RFLP, sequencing and SNPs analysis. Twenty-four of these clusters to genotype B, 13 to genotype C, and 19 to genotype D. High polymorphistic pattern was detected in the other 13 (18.6%) isolates, particularly at nt 25, 27, 43, 76, and 150, of which the genotype/subgenotype could not be determined by SNPs analysis. However, these 13 isolates showed consistent genotypes/subgenotypes classification by RFLP and phylogenetic analysis. The one remaining isolate (6809) was further analyzed by cloning and sequencing.

In addition, four isolates with mixed genotype B/C and C/D infection showed unidentified pattern by RFLP. These isolates are classified into genotype C and A by phylogenetic tree. Interestingly, specific SNPs pattern was detected in the 27 isolates of genotype A, E, G, and H: genotype A with A27 and A76; genotype E had A27, A43, C76, and T135; genotype G had A25, A27, A76 and T135; and genotype H had A76, and T150 (Supplementary Table 1).

3.3. Identification of mixed HBV genotypes.

Further analysis was done for seven isolates with inconsistent genotyping results between the RFLP and sequencing methods. Five isolates (7544, 998, 28159, Sur27, and 6809) were analyzed by cloning and sequencing. Of these, four isolates showed mixed HBV genotypes and one isolate showed single genotype. Isolate 998 had mixed genotype C/D infection, while isolates 28159, Sur27, and 6809 had genotype B/C infection, and all clones of sample 7544 had genotype A2. The other two isolates could not be further analyzed due to minimal serum samples.

Supplementary Table 1. This table shows SNPs at nine positions of preS2 sequences that could be used to distinguish genotype A through H as SNPs diagnostic sites for genotyping purposes.

Genotype	Position (nt)								
	20	25	27	43	45	76	96	135	150
A	G	G	A	G	G	A	C	C/T	C
B	A	T	A	G	C	A	A	C	C
C	G	G	T	G	G	C	C	C	C
D	A	G	A	A	G	A	C	T	T
E	G	G	A	A	G	C	C	T	C
F	G	G	T	G	C	A	C	C	T
G	G	A	A	G	G	A	C	T	C
H	G	G	T	G	G	A	C	C	T

4. Discussion

There is growing evidence that HBV genotype influences the course and outcome of the disease, and contributes to the outcome of treatment with interferons and pegylated interferons. HBV genotyping that was previously regarded as a research tool has been suggested as an important element to guide the choice of therapy by several investigators and national professional associations (Cooksley, 2010).

Recent data from multivariate analysis showed that HBV genotype was a strong predictor for sustained virological response in peginterferon treated-patients (Chan et al., 2003). Genotype information of chronic HBV infections enables practicing physicians to identify those patients at risk of disease progression and to determine the appropriate and optimal anti-viral therapy (Tanwar and Dusheiko, 2012). The Dutch guidelines state that the determination of genotype A or B is useful to predict a patient's likelihood of response to treatment with pegylated interferon (Buster et al., 2008). National guidelines from Germany and Sweden recommend checking the HBV genotype to aim a sustained success of pegylated interferon therapy (Cornberg et al., 2007; Lindh et al., 2008;). There is also a call from three regional bodies—AASLD, EASL and APASL— to widely assess the role of HBV genotype in clinical practice before adoption into the guidelines (EASL, 2009; Lok and McMahon, 2009; Liaw et al., 2012).

In the light of its usefulness in various epidemiological and clinical aspects, HBV genotype determination should be considered as important as other serological and virological parameters, particularly for HBV endemic regions, which are mostly of developing countries. A simple, reliable, rapid and low-cost assay is needed in laboratories with limited access to advanced technologies such as sequencing facilities. In this study, we developed an RFLP assay based on a short and highly polymorphic sequence of preS/S region of HBV genome.

In comparison with the conserved region of S gene, which is often used in RFLP as reported by some investigators, the high polymorphism within preS/S region has specific sites for some restriction enzymes (Figure 2 and 3), representing genotype information. Our analysis using simple software demonstrated that these restriction sites could distinguish eight HBV genotypes by their specific patterns (Figure 2 and 3). Supporting our results, phylogenetic tree analysis based on the preS/S region demonstrated an agreement in genotype clustering as those based on the complete genome tree (Figure 1a and 1b).

Genotyping results using the PCR-RFLP method showed 97% agreement with those based on sequencing and SNPs analysis. Our RFLP result was significantly verified with two other methods: nucleotide sequencing and SNPs analysis employing our previously described diagnostic sites (Nurainy et al., 2008; Thedja et al., 2011) that were based on the same preS/S region. Verification of RFLP results by phylogenetic tree clustering resulted in an agreement in 93.8% isolates. Furthermore, comparison and evaluation by all three methods successfully determined HBV genotype B, C, and D consistently in 80% isolates. This result confirms the validity of our RFLP and SNPs assays which is based on preS/S sequence for genotype B, C, and D. These findings support the previous study that PCR RFLP-based method should be considered as a good genotyping tool in addition to sequencing and hybridization (Ali et al., 2010). Thus, we suggest that the use of this simple RFLP genotyping method should be considered prior to interferon treatment, particularly in Asia and Pacific, where genotype B, C and D are predominant.

Additionally, this study showed the use of high variable sequence as genotyping tool based on SNPs diagnostic sites of preS/S region for genotype A, E, G, and H in addition to the previously described SNPs diagnostic sites for genotype B, C, and D (Nurainy et al., 2008; Thedja et al., 2011). We proposed these specific patterns as SNPs diagnostic sites for HBV/A, HBV/E, HBV/G, and HBV/H. Further study with larger number of isolates is needed for advance confirmation and elaboration.

In conclusion, this study demonstrates that HBV genotypic variation can be reflected by a short and partial polymorphic sequence of HBV genome, particularly within the preS2 sequence. Moreover, our findings strongly show that the particularly high polymorphism of HBV preS2 sequence can be assigned for genotyping purposes by using RFLP method based on the specific restriction site patterns. This simple and affordable HBV genotyping method is desirable and easily applicable in resource-limited settings such as the various developing countries in Asia and Pacific region, as well as a high throughput system for the analysis of large number of samples in epidemiological and clinical studies.

References

1. Ali, M.M., Hasan, F., Ahmad, S., and Al-Nakib, W., 2010. Comparative evaluation of INNO-LiPA HBV assay, direct DNA sequencing and subtractive PCR-RFLP for genotyping of clinical HBV isolates. Virology J 7, 111.
2. Aslam, M.A., Awan, F.R., Tauseef, I., Ali, S., Ahmad, N., Malik, N.A., Riaz, M.N., and Qureshi, J.A., 2008. Identification of hepatitis B virus core mutants by PCR-RFLP in chronic hepatitis B patients from Punjab, Pakistan. Arch Virol. 153(1), 163-170.
3. Buster, E.H.C.J., van Erpecum, K.J., Schalm, S.W., Zaaijer, H.L., Brouwer, J.T., Gelderblom, H.C., de Kneqt, R.J., Minke Bakker, C., Reesink, H.W., and Janssen, H.L., 2008. Treatment of chronic hepatitis B virus infection – Dutch national guidelines. Neth J Med 66, 292–305.
4. Chan, H.L.Y., Wong, M.L., Hui, A.Y., Chim, A.M.L., Tse, A.M.L., Hung, L.C.T., Chan, F.K.L., and Sung, J.J.Y., 2003. Hepatitis B virus genotype has no impact on hepatitis B e antigen seroconversion after lamivudine treatment. W J Gastroenterol 9, 2695-2697.
5. Chen, C.H., Lee, C.M., Hung, C.H., Wang, J.H., Hu, T.H., Changchien, C.S., and Lu, S.N., 2011. Hepatitis B virus genotype B results in better immediate, late and sustained responses to peginterferon-alfa in hepatitis-B-e-antigen-positive patients. J Gastroenterol Hepatol 26(3), 461-468.
6. Chu, C.J., and Lok, A.S.F., 2002. Clinical Significance of Hepatitis B Virus genotypes. Hepatology 35, 1274-1276.
7. Conberg, M., Protzer, U., Dollinger, M.M., Petersen, J., Wedermeyer, H., Ber, T., Jilg, W., Erhardt, A., Wirth, S., Schirmarcher, P., Fleig, W.E., Manns, M.P., 2007. Prophylaxis, diagnosis, and therapy of hepatitis B virus (HBV) infection: the German guidelines for the management of HBV infection. Z. Gastroenterol 45(12), 1281-1326.
8. Cooksley, .W.G.E., 2010. Do we need to determine viral genotype in treating chronic hepatitis B? J Viral Hepatitis 17, 601–610.
9. Ding, X., Mizokami, M., Yao, G., Xu, B., Orito, E., Ueda, R., and Nakanishi, M., 2001. Hepatitis B Virus Genotype Distribution among Chronic Hepatitis B Virus Carriers in Shanghai, China. Intervirology 44, 43-47.
10. Erhardt, A., Blondin, D., Hauck, K., Sagir, A., Kohnle, T., Heintges, T., Häussinger, D., 2005. Response to interferon alfa is hepatitis B virus genotype dependent: genotype A is more sensitive to interferon than genotype D. Gut 54, 1009–1013.
11. European Association for the Study of the Liver (EASL). 2009. EASL Clinical Practice Guidelines: Management of chronic hepatitis B. J Hepatol 50, 227–242.
12. Kao, J.H., Chen, P.J., Lai, M.Y., Chen, D.H., 2000. Hepatitis B genotypes correlate with clinical outcome in patients with chronic hepatitis B. Gastroenterol 118, 554-559.
13. Kramvis, A., Kew, M., Francois, G., 2005. Hepatitis B virus genotypes. Vaccine 23, 2409-2423.

14. Liaw, F.A., Kao, J.H., Piratvisuth, T., Chan, H.L.Y., Chien, R.N., Liu, C.H., Ed Gane, E., Locarnini, S., Lim, G., Han, K.H., Amarapurkar, D., Cooksley, G., Jafri, W., Mohamed, R., Hou, J.L., Chuang, W.L., Lesmana, L.A., Sollano, J.D., Suh, J.D., Omata, M., 2012. Asian-Pacific Consensus Statement On The Management Of Chronic Hepatitis B: A 2012 Update. Hepatol Int;DOI 10.1007/s12072-012-9365-4.

15. Lindh, M., J. E. Gonzalez, G. Norkrans, and Horal, P., 1998. Genotyping of hepatitis B virus by restriction pattern analysis of a pre-S amplicon. J. Virol. Methods 72, 163–174.

16. Lindh, M., Uhnoo, I., Bläckberg, J., Duberq, A.S., Friman, S., Fischler, B., Karlström, O., Norkrans, G., Reichard, O., Sangfeldt, P., Söderström, A., Weilan, O., Wejstål, R., Wiström, J., 2008. Treament of chronic hepatitis B infection: an update of Swedish recommendations. Scan J infect Dis 40 (6-7), 436-450.

17. Liu, C.J., Lo, S.C., Kao, J.H., Tsengi, P.T., Lai, M.Y., Ni, Y.H., Yeh, S.H., Chen, P.J., Chen, D.S., 2006. Transmission of occult hepatitis B virus by transfusion to adult and pediatric recipients in Taiwan. J Hepatol 44, 39-46.

18. Liu, W.C., Mizokami, M., Buti, M., Lindh, M., Young, K.C., Sun, K.T., Chi, Y.C., Li, H.H., Chang, T.T., 2006. Simultaneous quantification and genotyping of hepatitis B virus for genotypes A to G by Real-Time PCR and two-step melting curve analysis. J Clin Microbiol 44, 4491-4497.

19. Lok, .A.S.F., McMahon B.J., 2009. AASLD Practice Guideline Update. Chronic Hepatitis B: Update 2009. Hepatology 50, 1-35.

20. Mayerat, C., Mantegani, A., Frei, P.C. 1999. Does hepatitis B virus (HBV) genotype influence the clinical outcome of HBV infection? J Viral Hepat. 6(4), 299-304.

21. Mizokami, M., Nakano, T., Orito, E., Tanaka, Y., Sakugawa, H., Mukaide, M., Robertson, B.H., 1999. Hepatitis B virus genotype assignment using restriction fragment length polymorphism patterns. FEBS Lett. 450, 66–71.

22. Mulyanto, Depamede, S.N., Surayah, K., Tjahyono, A.A.H., Jirintai, Nagashima, S., Takahashi, M., Okamoto, H., 2010. Identification and characterization of novel hepatitis B virus subgenotype C10 in Nusa Tenggara, Indonesia. Arch Virol 155, 705-715.

23. Naito, H., Hayashi, S., Abe, K., 2001. Rapid and specific genotyping system for hepatitis B virus corresponding to six major genotypes by PCR using type-specific primers. J Clin Mirobiol 39, 362-364.

24. Nguyen, L.H., Ha, N.B., Vutien, P., Ha, NgB., Garcia, R.T., Trinh, H.N., Levitt, B.S., Nguyen, H.A., Nguyen, K.K., Keeffe, E.B., Nguyen, M.H., 2009. Prevalence of hepatitis B virus genotype B in Vietnamese patients with chronic hepatitis B. Hepatol Int 3, 461-467.

25. Norder, H., Couroucé, A.M., Coursaget, P., Echevarria, J.M., Lee, S.D., Mushahwar, I.K., Robertson, B.H., Locarnini, S., Magnius, L.O., 2004. Genetic diversity of hepatitis B virus strains derived worldwide: Genotypes, Subgenotypes, and HBsAg subtypes. Intervirology 47, 289-309.

26. Nurainy, N., Muljono, D.H., Sudoyo, H., Marzuki, S., 2008. Genetic study of hepatitis B virus in Indonesia reveals a new subgenotype of genotype B in east Nusa Tenggara. Arch Virol 153, 1057-1065.

27. Orito, E., Mizokami, M., Sakugawa, H., Michitaka, K., Ishikawa, K., Ichida, T., Okanoue, T., Yotsuyanagi, H., Iino, S., 2001. A Case-control study for clinical and molecular biological differences between Hepatitis B Viruses of genotypes B and C. Hepatology 33, 218-223.

28. Osiowy, C., Giles, E., 2003. Evaluation of the INNO-LiPA HBV genotyping assay for determination of Hepatitis B Virus genotype. J Clin Microbiol 41, 5473-5477.

29. Pas, S.D., Tran, N., de Man, R.A., Burghoorn-Maas, C., Vernet, G., and Niesters, H.G.J., 2008. Comparison of reverse hybridization, microarray, and sequence analysis for genotyping hepatitis B virus. J Clin Microbiol. 46(4), 1268-1273.

30. Sánchez-Tapias, J.M., Costa, J., Mas, A., Bruguera, M., Rodés, J., 2002. Influence of hepatitis B virus genotype on the long-term outcome of chronic hepatitis B in western patients. Gastroenterology 123(6), 1848-56.

31. Sumi, H., Yokosuka, O., Seki, N., Arai, M., Imazeki, F., Kurihara, T., Kanda, T., Fukai, K., Kato, M., Saisho, H., 2003. Influence of Hepatitis B Virus Genotypes on the Progression of Chronic Type B Liver Disease. Hepatology 37, 19-26.

32. Takahashi, K., Akahane, Y., Hino, K., Ohta, Y., Mishiro, S., 1998. Hepatitis B virus genomic sequence in the circulation of hepatocellular carcinoma patients: comparative analysis of 40 full-length isolates. Arch Virol 143, 2313-2326.

33. Tanwar, S., Dusheiko, G., 2012. Is there any value to hepatitis B virus genotype analysis ? Curr Gastroenterol Rep 14(1), 37-46.

34. Thedja, M.D., Muljono, D.H., Nurainy, N., Sukowati, C.H.C., Verhoef, J., Marzuki, S., 2011. Ethnogeographical structure of hepatitis B virus genotype distribution in Indonesia and discovery of a new subgenotype, B9. Arch Virol 156, 855-868.

35. Usuda, S., Okamoto, H., Iwanari, H., Baba, K., Tsuda, F., Miyakawa, Y., and Mayumi, M., 1999. Serological detection of hepatitis B virus genotypes by ELISA with monoclonal antibodies to type-specific epitopes in the preS2-region product. J Virol Methods 80, 97–112.

36. Usuda, S., Okamoto, H., Tanaka, T., Kidd-Ljunggren, K., Holland, P.V., Miyakawa, Y., and Mayumi, M., 2000. Differentiation of hepatitis B virus genotypes D and E by ELISA using monoclonal antibodies to epitopes on the preS2-region product. J Virol Method 87(1-2), 81-89.

37. Yuen, M.F., Sablon, E., Yuan, H.J., Wong, D.K.H., Hui, C.K., Wong, B.C.Y., Chan, A.O.O., and Lai, C.L., 2003. Significance of Hepatitis B Genotype in Acute Exacerbation, HBeAg Seroconversion, Cirrhosis-Related Complications, and Hepatocellular Carcinoma. Hepatology 37, 562-567.

38. Zeng, G., Wang, Z., Wen, S., Jiang, J., Wang, L., Cheng, J., Tan, D., Xiao, F., Ma, S., Li, W., Luo, K., Naoumov, N., V., Hou, J., 2005. Geographic distribution, virologic and clinical characteristics of hepatitis B virus genotype in China. J Viral Hepat 12, 609-617.

Figure 1 (a)

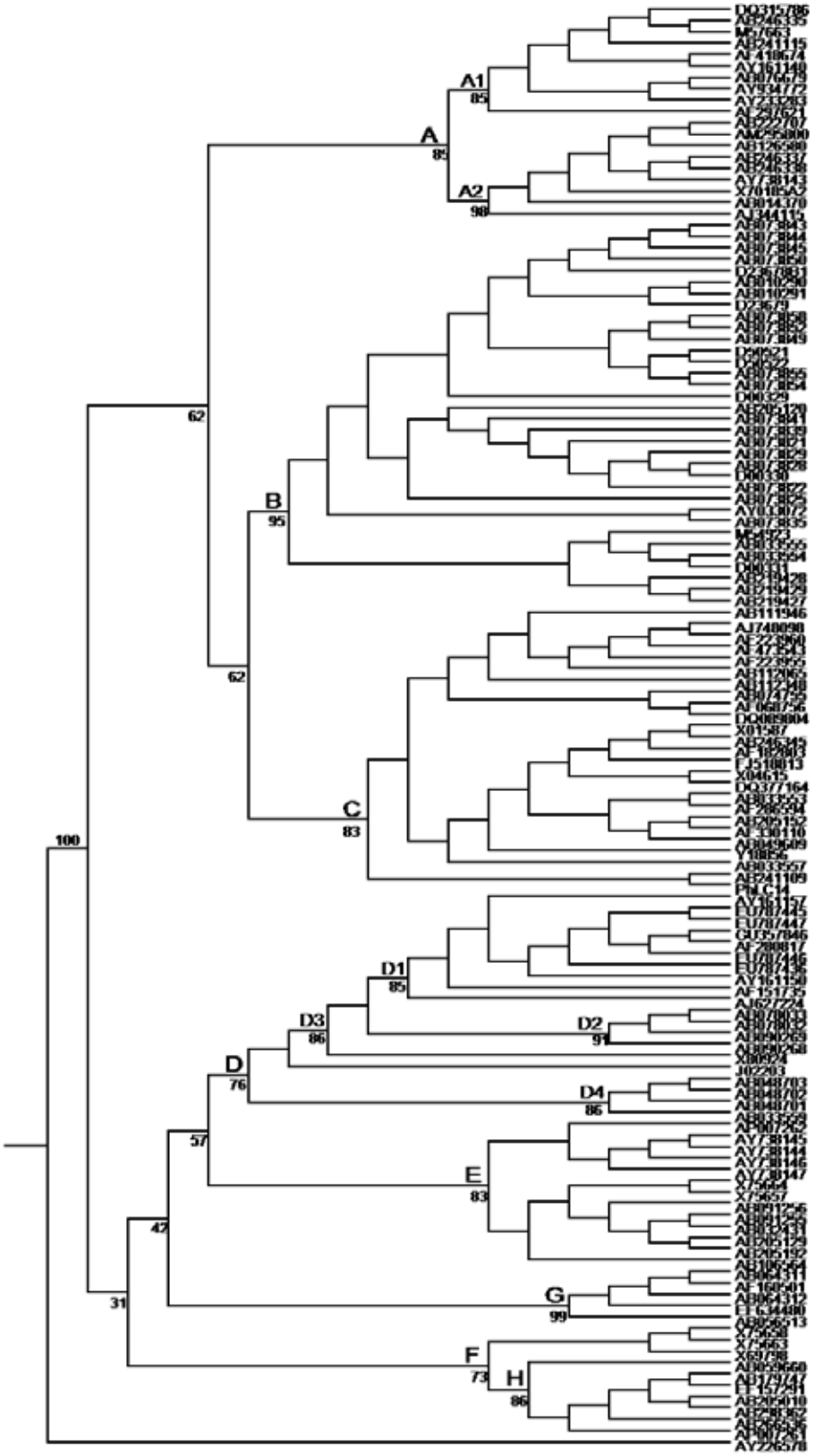


Figure 1 (b)

Figure 1 (a) & (b). Phylogenetic tree based on 126 complete genome sequence HBV isolates (a) and 126 preS/S sequence from the same isolates used for complete genome analysis (b). The tree was performed with neighbor joining algorithm based on the Kimura2-parameter distance estimation method. All the isolates were retrieved from the Genbank representing eight HBV genotypes with Wolly Monkey hepatitis B virus (AY226578) as outgroup. The clustering of HBV isolates in these two phylogenetic trees was consistent by both complete genome and preS/S sequences, with the exception one isolate belonged to genotype D4 (AJ627224) based on complete genome, however, they belonged to D1 based on PreS/S.

Figure 2

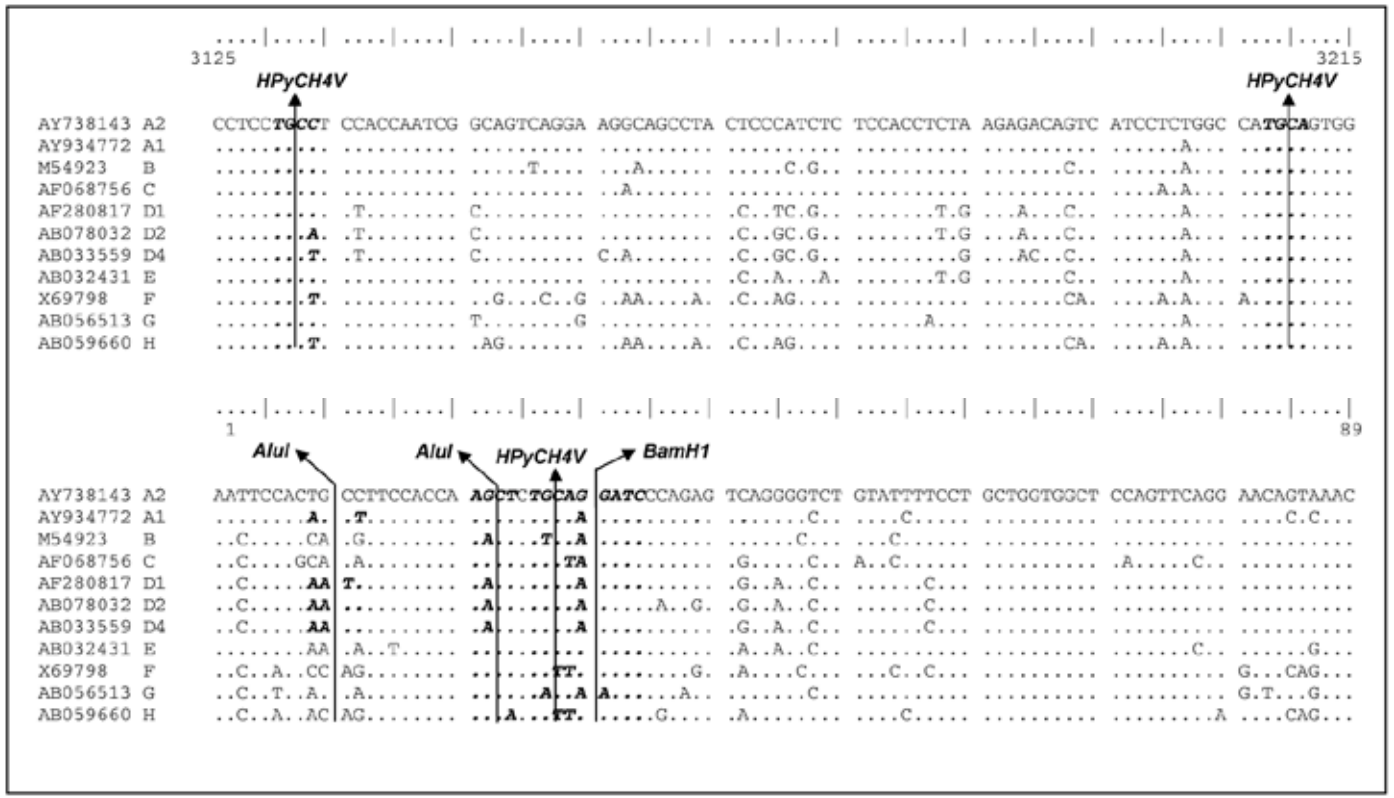
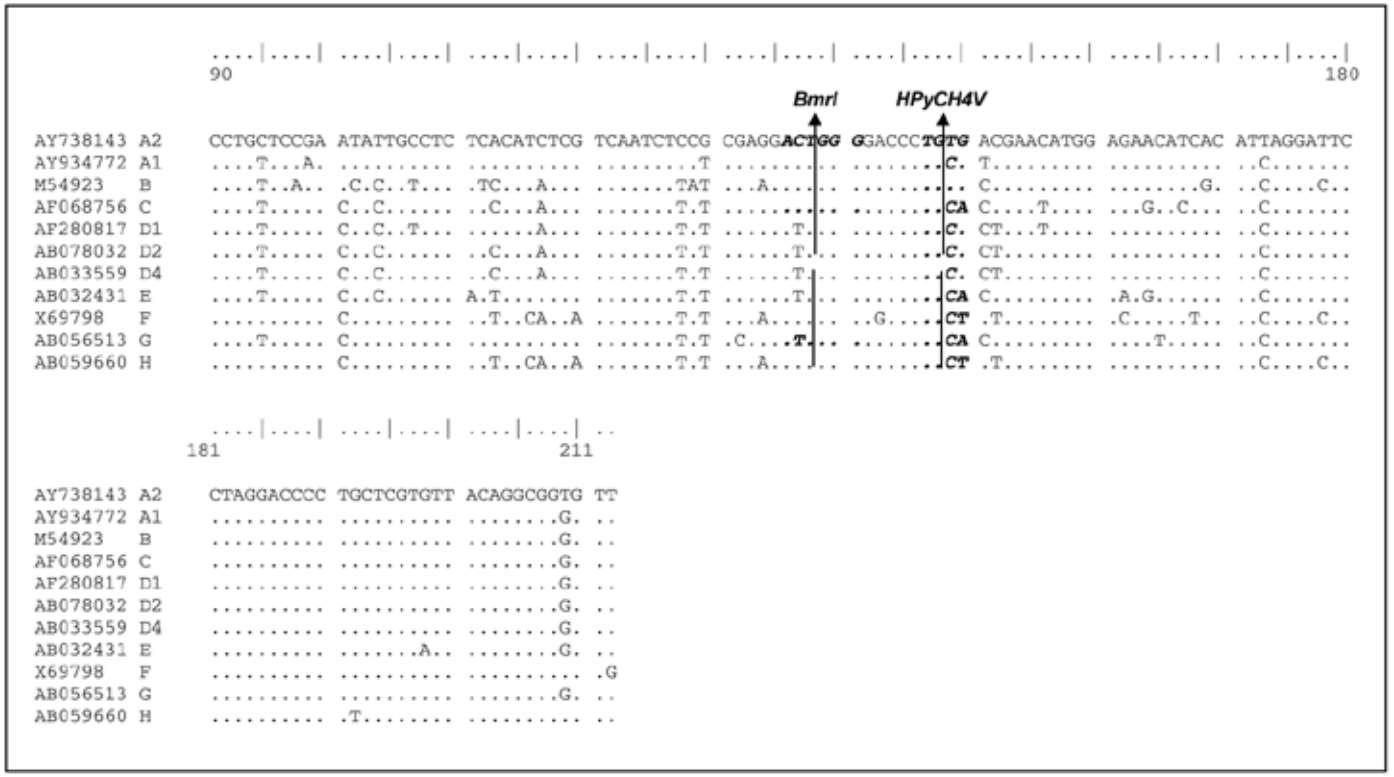


Figure 2. continued



Sequence alignment of the representative eight HBV genotypes/subgenotypes and their restriction sites of the four restriction enzymes within the preS/S sequences (302 bp) involving the highly polymorphic preS2 sequence. Four critical restriction sites were identified, *BamH1, HPyCH4V, Alu1*, and *Bmr1* resulting in different length bands. The numbering system was based on *EcoRI* site.

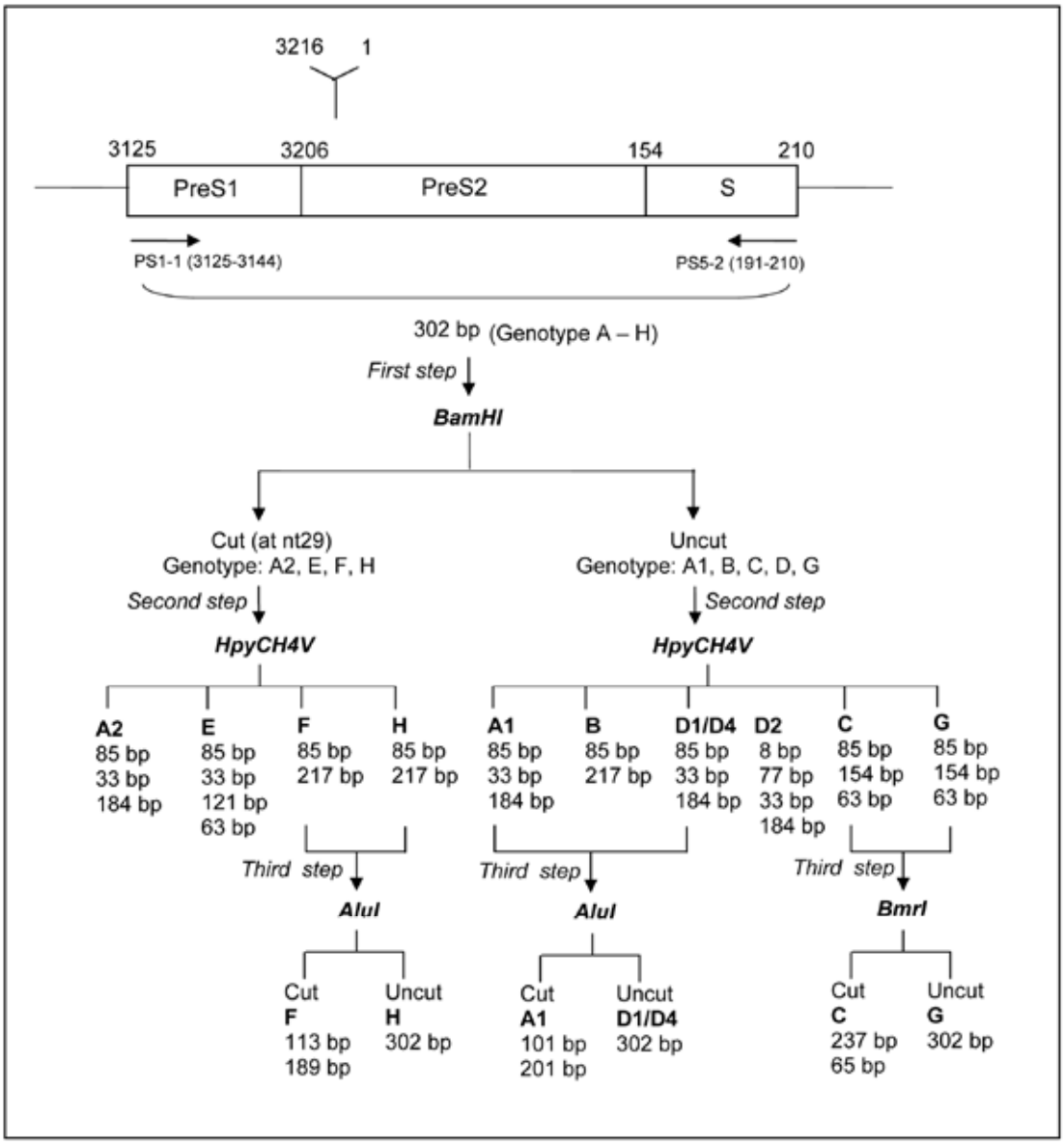


Figure 3. Schematic strategy of the PCR-RFLP based on preS/S sequence, nt3125-210 and the position of primers used in this study. There were three steps in identification of the eight genotypes, with the two first steps as major steps resulting into two major groups, the cut/digested group and uncut/undigested group. Different length fragments were observed for the eight genotypes.

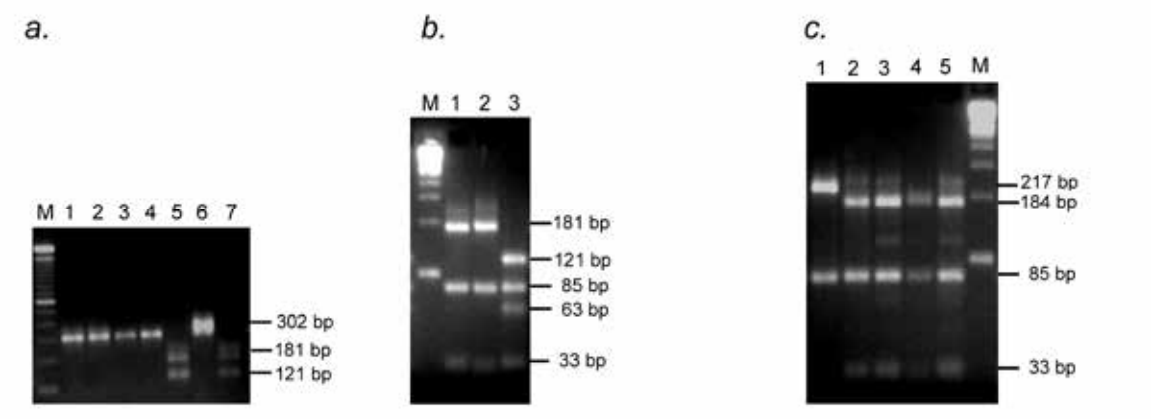


Figure 4. Series of RFLP patterns of HBV preS/S amplicons from the isolates studied. First digestion with BamH1 restriction enzyme revealed two major groups (a), the uncut/undigested (lane 1, 2, 3, 4, and 6) represented by a single 302 bp fragment and cut/digested amplicons (lane 5 and 7) represented by two fragments,181 bp and 121 bp. Second digestion with HPyCH4V obtained specific patterns in both cut/digested and uncut/undigested amplicons. Lane 1 and 2 of the cut/digested amplicons (b) showed three fragments, 184 bp, 85 bp, and 33 bp, that clustering these isolates into genotype A2. Four fragments, 121 bp, 85 bp, 63 bp, and 33 bp were detected in lane 3, clustering this isolate into genotype E. Lane 1 of the uncut/undigested amplicons (c) showed two fragments, 217 bp and 85 bp, clustering the isolates into genotype B. Lane 2, 3, 4, and 5 demonstrated 3 fragments, 184 bp, 85 bp, and 33 bp, classified these isolates to genotype A1 or D1/D4 that should be followed to the third step using Alu1. Lane M for 100 bp molecular weight ladder.

Figure 5. Phylogenetic tree of HBV isolates based on preS/S sequence of the isolates from the Netherlands and the Suriname. All the eight genotypes except genotype F were detected in all 100 isolates. HBV genotype A2, B, C, and D were found frequently compared than other genotypes.

CHAPTER 8

General Discussion, Future Perspectives, and Conclusion

General Discussion

After nearly 50 years since its discovery, hepatitis B virus (HBV) has been recognized as a major cause of chronic liver diseases with severe illnesses that gives rise to a global public health burden. Approximately 45% of the world's population, including those who live in many African and Asian countries, the Amazon Basin and parts of the Middle East, live in areas of high endemicity, where the prevalence of HBsAg is more than 8%, perinatal infection is the major mode of transmission with a lifetime risk of infection of more than 60%. The remainders of the world population live in countries of intermediate endemicity with HBsAg prevalence of 1-7%, and in areas of low endemicity where the prevalence is <1% [1,2].

In general, although HBV infection has been reduced by effective national infant vaccination programs implemented in 177 of 193 countries by 2008 [2,3], HBV infection continues to occur vertically through perinatal as well as horizontal transmission via blood and/or blood products or other body fluids in high-risk group populations. There are still around 50 million new cases diagnosed annually with 5–10% of adults and up to 90% of infants becoming chronically infected [4]. With an estimated 240 million people worldwide having chronic hepatitis B (CHB) [5], HBV infection appears as one of the most common causes of chronic liver failure, cirrhosis, fulminant hepatic failure and hepatocellular carcinoma.

CHB is a complex disease with intense interactions between viral, hepatocytes, immune response of host, and other factors that contribute to this interaction such as alcoholism, co-infection with other viruses, and life styles. Liver damage is mediated by the interaction of virus and host immune response; therefore, the dynamics of this interaction is of significance in determining the clinical outcome of the infection. Several factors have been identified as being responsible in this micro-circumstance of virus-host interaction, including genetic variety of both the virus and the host. In addition to the knowledge of host genetic variety, understanding of HBV variety would be necessary in the development of prevention and treatment strategies for HBV infection worldwide.

HBV is a DNA virus with high genome variability as the consequence of reverse transcription in its replication that unfortunately does not possess proof-reading capacity, resulting in high nucleotide diversity. Nucleotide divergences over the HBV entire genome have led to the identification of several genotypes, currently designated to genotypes A, B, C, D, E, F, G, and H, with two additional newly proposed genotypes, I and J [6,7].

With some genotypes having several subgenotypes, eight of nine subgenotypes of genotype B (B1, B2, B3, B4, B5, B7, B8, and B9) and seven of 16 subgenotypes of genotype C (C1, C2, C5, C7, C8, C9, and C10) have been identified in East Asia and Southeast Asia. Further, indigenous populations of Japan that moved to Arctic region including Northern Canada, Alaska, and Greenland retained the HBV/B6 subgenotype [8]. Of all existing HBV genotypes, B (with nine subgenotypes) and C (with seven subgenotypes) are specific for Asia, which is home to the Austronesian-speaking populations.

A previous comprehensive study (**Chapter 2**) clearly showed the genetic variety of HBV in heterogeneous populations with well-defined ethnic background in Indonesia. HBV genetic diversity distinctly showed diversification across the Indonesian archipelago, with subgenotype B3 dominant in the western islands, five subgenotypes of genotype B (B3, B5, B7, B8, and B9) and two subgenotypes of genotype C (C1 and C2) in the eastern islands, and C6 majoring in Papua, the east-most island of Indonesia. Interestingly, C6 was restricted in the Papuan-speaking populations. The other unique finding in the study was the observation that B2, which is specific for Chinese populations of mainland Asia and Taiwan, was dominant in Indonesian of Chinese ethnic origin.

The phenomenon that HBV genetic characteristics are retained in certain population groups was also demonstrated by another study on the characteristics of HBV genotype C (HBV/C) in the Asia and Pacific region (**Chapter 3**). Strains of HBV/C in Austronesian speaking-populations in the East to Southeast Asia were different not only in the genetic pattern of the virus, but also in the immunoepitope characteristics, from those Papuan speaking-populations in the Papua, Papua New Guinea (PNG) to Pacific islands. Based on these differences, HBV/C strains can be divided into two types, the Asian and the Papua-Pacific. Another important finding from that study, is that the distribution of HBV genetic characteristics, particularly for genotype B and C, emerges in a specific pattern from the East and Southeast Asia in the west through the Pacific in the east, forming a west (subtype *adrq*+) to east (subtype *adrq*-) gradient. Uniquely, the gradient change occurs in Papua and Papua New Guinea (PNG), where some interfaces of subtype *adrq,* namely *adrq*-interderminate A159/A177 and V159/V177, were present (**Chapter 3**). We suggest that Papua and PNG is the admixture region of HBV/C genomic evolution from HBV/C Asian type to HBV/C Papua-Pacific type based on subtype *adr* characteristics.

The findings in these two studies provided the first direct evidence that HBV genetic diversity is indeed associated with ethnic and geographical origin of host. Definitely, further studies in more heterogeneous populations are needed to understand the significance of the HBV genome diversity and host-ethnic variety in regard to the clinical

and public health implication that challenge health and medical professionals in the control of hepatitis B.

Further, the information and analysis of association between HBV genome diversity and ethnogeographical of host could be integrated into archeological as well as linguistic and genetic perspectives in tracing and understanding the entry and subsequent human migrations of the modern humans. Therefore, the genetic diversity of microorganism of infectious diseases, ethnic variety of host, and geographical origin should be considered partly to contribute to the genetic, linguistic and archeological perspectives.

Significance of HBV Genetic Diversity to Public Health and Clinical

Public health issues associated with HBV genetic diversity have been investigated since more than 10 years, and recently, the clinical consequences as well. Nucleotide substitutions or amino acid mutations in surface gene or HBsAg, respectively, have been shown associated with vaccine failure [9] and undetectable of HBsAg in regular blood donors (**Chapter 4**). Our previous study showed that mutations in HBsAg reduced the antigenicity and conformation of the surface protein (**Chapter 5**) that resulted in HBsAg undetectability albeit in the presence of HBV DNA. The undetectable HBsAg in the presence of HBV DNA or occult hepatitis B should be considered as a hidden source of hepatitis B that could be transmitted to recipients. This situation emphasizes the risk, public health and clinical messages of HBsAg mutations. The epidemiology of occult hepatitis B needs to be further determined and characterized; a strategy to screen and manage those with HBsAg-negative and anti-HBc positive is needed.

Interaction between virus and immune system of the host results in different clinical outcome that forms a specific natural course of disease. Genetic variety of virus and host indeed influences the interaction and the clinical outcome. Apart from host factors, it is important to identify relevant viral factors involved in the progression of HBV-related liver diseases. HBV mutants have been shown to affect the natural course and treatment outcomes of chronic. For example, mutations in the precore/BCP regions have been associated with the progression of liver disease and higher incidence of hepatocellular carcinoma (HCC) [10]. The predominant mutation of the precore region involves a G-to-A change at nucleotide 1896 (G1896A), which creates a premature stop codon at codon 28. The critical mutation of the BCP region (nucleotides 1742-1849) occurs with an adenine (A) to thymine (T) transversion at nucleotide 1762 together with a guanine (G) to adenine (A) transversion at nucleotide 1764 (A1762T/G1764A) [11,12,13].

In a study on CHB patients (**Chapter 6**), a relationship between the genetic characteristics of the virus and the HBeAg status of the patients was observed, with a higher rate of HBeAg positive in patients infected with genotype B than those with genotype C. This is consistent with previous studies that reported the higher rate of HBeAg seroconversion by genotype B than by genotype C [17,19]. This study also found that HBsAg level, which is a surrogate marker of cccDNA translational activity [16], was highest in the IT group, and lowest LR group. In can be explained that in the IT phase, no immune reaction exists against HBV and its antigens can be maintained in a host. However, vigorous immune reaction occurs in the IC phase as marked by increased ALT levels. Following the success of the host immunity, the patient enters into the 'inactive phase' (LR) of HBV infection. The presence of HBsAg in all phases, despite its different levels in each phase, suggests that HBV cccDNA transcriptional activity occurs throughout the natural history of CHB.

The study observed that the HBsAg level was correlated with the HBV DNA level in overall patients regardless of the phase. Analysis by each phase showed that this correlation was strong in the IT and IC, none in the LR, and low in the ENH groups. Based on this finding, HBsAg synthesis seems to occur parallel with HBV synthesis during early phases. But after the immune clearance by the host, this connection appeared to be slowly dissociated as the phase progressed. A possible explanation for this finding is that HBsAg synthesis has a distinct pathway from HBV DNA synthesis, and it may be under the influence of different immune control mechanisms. Meanwhile, the BCP T1762/A1764 mutations were identified similarly in all phases of CHB. It may indicate that the deleterious effect of these mutations may take place any time in CHB patients. Taken together, this study emphasizes the need to regularly monitor CHB patients.

The Importance of HBV Genotype Information

In addition to the epidemiological importance and its association with anthropology and the past human migration in peopling of the world, HBV genotype has been reported to influence the course of disease. Many studies have shown that genotype C is associated with more severe liver disease, higher occurrence of HCC, and related to later HBe antigen seroconversion than genotype B [17,18,19,20]. Meanwhile, genotype D is associated with severe liver disease and HCC than genotype A [21]. There are multiple subgenotypes that may also have clinical significance. For example, B1 (Bj for Japan) occurs in Japan, whereas B2 (Ba for Asia) occurs in Taiwan, China and Vietnam and appears to be a recombination of genotype B and C [8] and has a different natural history from genotype B1.

HBV genotype has also been associated with outcome to interferon treatment. Studies conducted from 1993 to 2007 on patients treated with conventional interferon in Asia and Europe showed higher sustained response in genotype B than in genotype C [21,22], and in genotype A than in genotype D [23]. Similar pictures have been obtained from studies on patients treated with pegylated interferon, which is slowly replacing conventional interferon worldwide. In CHB patients with genotypes A, B, C and D, the response rate was highest in genotype A, intermediate in genotypes B and D, and lowest in genotype D [24].

Taking these lines of evidence together, the HBV genotype is currently considered as important information in addition to serological profile and molecular characteristics [14,17,26]. HBV genotyping previously regarded as a research tool, is now considered by several investigators and professional associations as one of the items to guide the choice of therapy [27]. The Dutch guidelines state that the determination of genotype A or B is useful to predict a patient's likelihood of response to treatment with pegylated interferon [28]. The German guidelines recommend checking the genotype to aim a sustained success of therapy [29]. The Swedish national guidelines (2008) also recommend HBV genotype being ascertained prior to pegylated interferon treatment [30]. There is also a call from three regional bodies—AASLD, EASL and APASL— to widely assess the role of HBV genotype in clinical practice before its inclusion into the guidelines [1, 3, 32,33].

HBV genotyping by phylogenetic analysis based on nucleotide sequences has been used as a gold standard. However, this is not an appropriate method for large scale genotyping. In our study, specific Single Nucleotide Polymorphisms (SNPs) were discovered in the highly variable preS2 sequence that could be used as a simple genotyping tool by RFLP (**Chapter 3**), in addition to the diagnostic sites previously reported from our center [34]. By employing the identified SNPs, we have developed a simple HBV genotyping system using RFLP method that can classify genotypes A through H (**Chapter 7**). This new RFLP system is simple and useful for rapid and sensitive HBV genotyping in centers that do not have sequencing facilities, when either epidemiological or transmission studies are carried out in large scale. It can be performed for clinical research, with implications for improved individualized treatments of HBV patients.

Future perspectives

The outcome of HBV infection depends on the interaction between viral and host immune system. HBV has high genetic diversity compared to other DNA viruses. On the other side, Indonesia has hundreds of ethnic groups inhabiting the major and in the hundred small islands across the archipelago. Distribution of HBV genetic diversity is ethnically and geographically specific in these various Indonesian ethnic populations. In addition to the information of HBV characteristics in homogenous ethnic populations, a further study on how this HBV high genetic diversity contributes to the clinical outcome of disease in a very diverse population such in Indonesia is needed. Ideally, these investigations should be population-based studies that compare persons with different genotypes, conducted prospectively over a long period of time to precisely identify the risk of sequel associated with each genotype and subgenotype, such as have been done in Asia and Alaska [35,36,37].

Mutations that cause HBsAg undetectability in blood donors or occult hepatitis B (defined as infection with detectable HBV DNA and undetectable HBsAg in patients' blood) may spread to recipients, and therefore, a threat for public health. A larger study should be performed to understand the magnitude of the problem in blood donors, and to develop a national strategy on blood donor screening policy. In addition, the knowledge on HBV molecular characteristics in occult hepatitis B in Indonesian blood donors is required as the basic information in understanding pathogenicity of occult hepatitis B in asymptomatic carriers and in providing data for development of hepatitis B diagnostic tools. HBV genome variety in various ethnic populations together with HBV molecular characteristics in occult hepatitis B in asymptomatic Indonesian blood donors should be integrated and further studied for a more potent hepatitis B vaccine development.

Conclusion

Distribution of HBV genetic diversity in Indonesia is associated and follows the ethnic background and geography of host. Internal viral genomic evolution including recombination, and pressure of host immune response are suggested to give strong contribution in the genetic diversity of HBV. Thus, HBV genetic diversity is partly due to virus-host interactions and partly due to parallel evolution in geographically distinct areas. Some of these variations have significant impact to public health and clinical settings. Mutations within HBsAg, particularly in the *'a'* determinant, may cause HBsAg undetectability or occult hepatitis B using commercial diagnostic kit and this situation

could mislead diagnosis and disease management that give significant impact to public health particularly in blood donors and recipients. The findings of this study alarm us to alert for the spread of HBV infection due to unrecognized infection, and thus, the national blood donor policies for blood donor candidates, based on HBsAg detection, should be reviewed. Furthermore, this situation should be a warning to clinicians to be more careful in giving blood and or blood products transfusion. In clinical perspective, HBV genetic variety has been demonstrated to be associated with HBeAg seroconversion, which is higher in genotype B than in genotype C. This information is important for the clinical management of CHB patients in Indonesia.

We conclude that HBV genetic diversity gives major information on understanding the molecular pathogenesis of hepatitis B that may contribute to the efforts to reduce the morbidity and mortality. This knowledge is also important to understand the molecular epidemiology of HBV infection in a hepatitis B endemic country such as Indonesia. This HBV diversity data can be used as important information in diagnostic kits and vaccine developments that is ethnically specific. Further, data of HBV genomic variety should be linked to non-clinical aspects such as linguistic, genetic and archeological perspectives in order to maximize human health. This series of studies deliver broad information from basic aspects of the HBV genetic diversity to the public health and clinical implications, and introduce the development of genotyping tool useful particularly for analysis of large samples in epidemiological and clinical studies.

References

1. Lavanchy, D. Hepatitis B virus epidemiology, disease burden, treatment, and current and emerging prevention and control measures. J. Viral. Hepat. 2004;11: 97–107.
2. World Health Organization. Hepatitis B vaccines. Weekly Epidemiological Record. 2009; 40: 405-420.
3. WHO IVB database. Hepatitis B - 12 April 2011 update. http://www.who.int/immunization/topics/hepatitis_b/en/index.html.
4. Merican, I., Guan, R., Amarapuka, D., Alexander, M.J., Chutaputti, A., Chien, R.N., *et al*. Chronic hepatitis B virus infection in Asian countries. J. Gastroenterol. Hepatol. 2000; 15(12): 1356-1361.
5. Lai, C.L., Dienstag, J., Schiff, E., Leung, N.W.Y., Atkins, M., Hunt, C., *et al*. Revalence and clinical correlates of YMDD variants during lamivudine therapy for patients with chronic hepatitis B. Clin. Infect. Dis. 2003; 36: 687–696.
6. Huy, T.T.T., Ngoc, T.T., Abe, K. New complex recombination genotype of hepatitis B virus identified in Vietnam. J. Virol. 2008; 82: 5657–5663.

7. Tatematsu, K., Tanaka, Y., Kurbanov, F., Sugauchi, F., Mano, S., Maeshiro, T., *et al*. A genetic variant of hepatitis B virus divergent from known human and ape genotypes isolated from a Japanese patient and provisionally assigned to new genotype J. J. Virol. 2009; 83(20): 10538–10547.

8. Sakamoto, T., Tanaka, Y., Simonetti, J., Osiowy, C., Berresen, M.L., Koch, A., *et al*. Classification of hepatitis B virus genotype into 2 major types based on characterization of a novel subgenotype in Arctic indigenous populations. J. Infect. Dis. 2007; 196:1487–1492.

9. Ho, M.S., Mau, Y.C., Lu, C.F., Huang, S.F., Hsu, L.C., Lin, S.R., *et al*. Patterns of circulating hepatitis B surface antigen variants among vaccinated children born to hepatitis B surface antigen carrier and non-carrier mothers: a population-based comparative study. J. Biomed. Sci. 1998; 5: 355–362.

10. Tong, M.J., Blatt, L.M., Kao, J.H., Cheng, J.T., Corey, W.G. Basal core promoter T1762/A1764 and precoreA1896 gene mutations in hepatitis B surface antigen-positive hepatocellular carcinoma: a comparison with chronic carriers. Liver. Int. 2007; 27:1356-1363.

11. Ren, X., Xu, Z., Liu, Y., Li, X., Bai, S., Ding, N., *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. J. Viral. Hepat. 2010; 17: 887-895.

12. Chan, H.L.Y. Significance of hepatitis B virus genotypes and mutations in the development of hepatocellular carcinoma in Asia. J Gastroenterol Hepatol. 2011; 26: 8-12.

13. Zheng, J.X., Zeng, Z., Zheng, Y.Y., Yin, S.J., Zhang, D.Y., Yu, Y.Y., *et al*. Role of hepatitis B virus base core and precore/core promoter mutations on hepatocellular carcinoma in untreated older genotype C Chinese patients. J. Viral. Hepatitis. 2011;18:1365-2893 (e423–e431).

14. Orito, E., Mizokami, M., Sakugawa, H., Michitaka, K., Ishikawa, K., Ichida, T., *et al*. A Case-control study for clinical and molecular biological differences between Hepatitis B Viruses of genotypes B and C. Hepatology. 2001; 33: 218-223.

15. Zeng, G., Wang, Z., Wen, S., Jiang, J., Wang, L., Cheng, J., *et al*. Geographic distribution, virologic and clinical characteristics of hepatitis B virus genotype in China. J. Viral. Hepat. 2005; 12: 609-617.

16. Chan, H.L.Y., Wong, V.W.S., Tse, A.M.L., Tse, C.H., Chim, A.M.L., Chan, H.Y., *et al*. Serum Hepatitis B Surface Antigen Quantitation can Reflect Hepatitis B Virus in the liver and predict treatment Response. Clin. Gastoentero. Hepatol. 2007;5:1462-1468.

17. Kao, J.H., Chen, P.J., Lai, M.Y., Chen, D.S. Hepatitis B genotypes correlate with clinical outcomes in patients with chronic hepatitis B. Gastroenterology. 2000; 118: 554-559.

18. Sakugawa, H., Nakasone, H., Nakayoshi, T., Orito, E., Mizokami, M., Yamashiro, T. Preponderance of hepatitis B virus genotype B contributes to a better prognosis of chronic HBV infection in Okinawa, Japan. J. Med. Virol. 2002; 67: 484–489.

19. Chan, H.L., Wong, M.L., Hui, A.Y., Hung, L.C., Chan, F.K., Sung, J.J. Hepatitis B virus genotype C takes a more aggressive disease course than hepatitis B virus

genotype B in hepatitis Be antigen-positive patients. J. Clin. Microbiol. 2003; 41: 1277-1279.

20. Pujol, F.H., Navas, M.C., Hainaut, P., Chemin, I. Worldwide genetic diversity of HBV genotypes and risk of hepatocellular carcinoma. Cancer. Lett. 2009; 286: 80-88.

21. Thukar, V., Guptan, R.C., Kazim, S.N., Malhotra, V., Sarin, S.K. Profile, spectrum and significance of HBV genotypes in chronic liver disease in the Indian subcontinent. J. Gastroenterol. Hepatol. 2002; 17: 165–170.

22. Craxi, A., Di Bona, D., Camma, C. Interferon alpha for HBeAg positive chronic hepatitis B. J. Hepatol. 2003; 39(Suppl.): S99–S105.

23. Erhardt, A., Blondin, D., Hauck, K. Response to interferon alpha is hepatitis B virus genotype dependent: genotype A is more sensitive to interferon than genotype D. Gut. 2005; 54: 1009–1013.

24. Weigand, J., Hasenclever, D., Tillmann, H.L. Should treatment of hepatitis B depend on hepatitis B virus genotypes? A hypothesis generated from an explorative analysis of published evidence. Antivir. Ther. 2008; 13: 211–220.

25. Flink, H.J., van Zonneveld, M., Hansen, B.E., de Man, R.A., Schalm, S.W., Janssen, H.L. Treatment with Peg-interferon alfa 2b for HBe antigen positive chronic hepatitis B: HBsAg loss is associated with HBV genotype. Am. J. Gastroenterol. 2006;101: 297–303.

26. Nguyen, L.H., Ha, N.B., Vutien, P., Ha, NgB., Garcia, R.T., Trinh, H.N., *et al*. Prevalence of hepatitis B virus genotype B in Vietnamese patients with chronic hepatitis B. Hepatol. Int. 2009; 3: 461-467.

27. Cooksley, W.G.E. Do we need to determine viral genotype in treating chronic hepatitis B? J. Viral. Hepat., 2010, 17, 601–610.

28. Buster, E.H.C.J., van Erpecum, K.J., Schalm, S.W., Zaaijer, H.L., Brouwer, J.T., Gelderblom, H.C., *et al*. Treatment of chronic hepatitis B virus infection – Dutch national guidelines. Neth. J. Med. 2008; 66:292–305.

29. Cornberg, M., Protzer, U., Dollinger, M.M., Petersen, J., Wedemeyer, H., Berg, T., *et al.* The German guideline for the management of hepatitis B virus infection: short version. J. Viral. Hepat. 2008; 15(Suppl. 1): 1–21.

30. Lindh, M., Uhnoo, I., Bläckberg, J., Duberg, A.S., Sriman, S., Fischler, B., *et al*. Treatment of chronic hepatitis B infection: an update of Swedish recommendations. Scand. J. Infect. Dis. 2008; 40: 436–450.

31. Lok, A.S.F., McMahon, B.J. AASLD Practice Guideline Update. Chronic Hepatitis B: Update 2009. Hepatology. 2009; 50: 1-35.

32. European Association for the Study of the Liver (EASL). EASL Clinical Practice Guidelines: Management of chronic hepatitis B. J. Hepatol. 2009; 50: 227–242.

33. Liaw, F.A., Kao, J.H., Piratvisuth, T., Chan, H.L.Y., Chien, R.N., Liu, C.H., *et al*. 2012. Asian-Pacific Consensus Statement On The Management Of Chronic Hepatitis B: A 2012 Update. Hepatol. Int.;DOI 10.1007/s12072-012-9365-4.

34. Nurainy, N., Muljono, D.H., Sudoyo, H., Marzuki, S. Genetic study of hepatitis B virus in Indonesia reveals a new subgenotype of genotype B in east Nusa

Tenggara. Arch. Virol. 2008; 153:1057-1065.

35. Watanabe, K., Takahashi, T., Takahashi, S., Okoshi, S., Ichida, T., Aoyagi, T. Comparative study of genotype B and C hepatitis B virus-induced chronic hepatitis in relation to the basic core promoter and precore mutations. J. Gastroenterol. Hepatol. 2005; 20(3): 441-449.

36. Yu, M.W., Yeh, S.H., Chen, P.J., Liaw, Y.F., Lin, C.L., Liu, C.J., *et al.* Hepatitis B Virus Genotype and DNA Level and Hepatocellular Carcinoma: A Prospective Study in Men. J. Natl. Cancer. Inst. 2005; 97: 265-272.

37. Livingston, S.E., Simonetti, J.P., McMahon, B.J., Bulkow, L.R., Hurlburt H.J., Homan, C.E., *et al*. Hepatitis B Virus Genotypes in Alaska Native People with Hepatocellular Carcinoma: Preponderance of Genotype F. J. Infect. Dis. 2007; 195: 5-11.

Samenvatting

Samenvatting

Hepatitis B virus (HBV) infectie is een serieus probleem in de publieke gezondheid door de complicaties die het gevolg zijn van chronische hepatitis B (CHB). Deze complicaties variëren van asymptomatische dragerschap tot levercirrose en leverkanker. CHB heeft een significante impact op morbiditeit en mortaliteit, en heeft een substantiële invloed op de economie vooral in Oost en Zuidoost Azië en delen van de Pacific waar ongeveer 75% HBV van de chronische dragers leven. Het begrijpen van de genetische variëteit van HBV en de genetische achtergrond van de gastheer is noodzakelijk voor de behandeling en voor de ontwikkeling van preventieve maatregelen.

HBV is een gedeeltelijk dubbelstrengs DNA virus met vier Open Reading Frames (ORF), dat zijn de oppervlakte (O), kern (K), polymerase (G) en X genen, die elkaar overlappen. Deze overlapping van de ORF's in het genoom, de hoge mutatie snelheid en het ontbreken van "proofreading", gecombineerd met druk van het immuunsysteem van de gastheer leidt tot een hoge variatie binnen het HBV genoom. Deze variatie heeft grote consequenties voor de volksgezondheid en voor de patiënt zelf. Dit benadrukt het belang van kennis en begrip van het HBV genoom om te komen tot betere behandelen en zorg voor patiënten en dus voor het verlagen van de mortaliteit en morbiditeit.

De hoge variatie in het genoom van HBV heeft geleid tot de classificatie van een aantal genotypen, genotype A tot H. Tevens zijn twee nieuwe genotypen sinds kort geïdentificeerd in Azië: genotype I en genotype J. Interessant is de geografische en etnische verspreiding van de HBV genotypes, in het bijzonder bij de Indonesische bevolkingsgroepen zoals beschreven in **Hoofdstuk 2** en in de regio Zuid Oost Azië en de Pacific zoals beschreven in **Hoofdstuk 3**. In **Hoofdstuk 2**, laat de HBV genetische diversiteit een duidelijk geografisch patroon binnen de Indonesische archipel zien met het subgenotype B3 dominant op de westerse eilanden: vijf subgenotypen van het genotype B (B3, B5, B7, B8 en B9) en twee subgenotypen van genotype C (C1 en C2) zijn aanwezig op de oostelijke eilanden en het genotype C6 bij de meerderheid in Papoea, het meest oostelijk gelegen eiland van Indonesië. Interessant is dat C6 beperkt is tot de Papoea sprekende bevolking.

Een andere unieke ontdekking in de studie was de observatie dat genotype B2, dat specifiek thuis hoort bij Chinese bevolkingsgroepen op het vaste land van Azië en Taiwan, ook dominant is in de Indonesische bevolking van Chinese etnische oorsprong.

Het fenomeen dat HBV genetische karakteristieken blijven voort bestaan in bepaalde bevolkingsgroepen, is ook geconstateerd in een andere studie betreffende de

karakteristieken van het HBV genotype C (HBV/C) zoals beschreven in **Hoofdstuk 3**. Isolaten van HBV/C bij Austronesian sprekende bevolking in het Oosten van Zuidoost Azië waren verschillend van de isolaten van de inwoners van Papoea en Papoea-Nieuw-Guinea (PNG). Niet alleen het genetische patroon van het virus, maar immuun-epitopen waren verschillend. Dus, gebaseerd op de moleculaire karakteristieken, immuun-epitopen, geografisch en etnische origine van de gastheer kunnen HBV/C isolaten onderverdeeld worden in twee types, het Aziatisch en het Papoea/Pacific type C. Een andere belangrijke bevinding van deze studie is dat de verspreiding van de HBV genotypen zich verspreiden volgens een bepaald patroon: in Oost en Zuidoost Azië met in het Westen subtype *adrq+* en naar het Oosten het subtype *adrq-* in toenemende prevalentie. Uitzonderlijk, is de verandering die zich voordoet in Papoea en PNG, waar enkele variaties van het subtype *adrq*, met name *adrq-*interderminate A159/A177 aanwezig is. Dit suggereert dat Papoea en PNG de regio is de oorsprong is van de ontwikkeling van de HBV/C genetische variatie: het HBV/C Aziatisch type en het HBV/C Papoea /Pacific type: gebaseerd op de karakteristieken van het subtype *adr*.

De grote genoom diversiteit van HBV heeft grote implicaties voor de volksgezondheid. De aanwezigheid van occulte hepatitis B bij bloeddonoren is mogelijk te wijten aan mutaties in het oppervlakte eiwit HBsAg (**Hoofdstuk 4**). Mutaties in het HBsAg verlagen de antigeniciteit (**Hoofdstuk 5**). Het resultaat is dat het HBsAg niet te detecteren is, ondanks het feit dat HBV DNA wel aanwezig is. Dit is een ernstig risico voor de volksgezondheid. De epidemiologie van het occulte Hepatitis B moeten verder uitgewerkt en gekarakteriseerd worden en er moet een nieuw beleid komen voor deze bedreigende situatie.

Mutaties in het HBV, precore - en basal core promotor-gen (BCP) zijn geassocieerd met progressieve leverziekte en een hogere incidentie van hepatocellulair carcinoom (HCC). De meest voorkomende mutatie van in het precore gen is de "G-to-A" mutatie van nucleotide 1896 (G1896A) en een "G-to-A" mutatie op nucleotide 1764.

In een studie met CHB patiënten (**Hoofdstuk 6**) hebben wij gevonden dat de precore A 1896 mutatie vaker aanwezig was bij patiënten in latere fases van CHB infecties en bij oudere patiënten.. De BCP T1762/A1764 mutaties werden echter gevonden in alle fases van CHB. Deze ontdekking geeft aan dat het schadelijke effect van deze mutaties ieder moment kan plaatsvinden bij CHB patiënten. BCP mutaties komen het meest voor bij genotype C/subtype *adr* en minder bij het genotype B/subtype *adw*. In dit onderzoek is ook waargenomen dat HBsAg niveaus vergelijkbaar waren over alle CHB fases. HBsAg niveaus correleerden niet met de HBV DNA niveaus in het natuurlijke beloop van een CHB. Dit suggereert dat HBV cccDNA transcriptie activiteit plaatsvindt in alle fases in

het natuurlijk verloop van CHB. Deze studie benadrukt de noodzaak om patiënten met CHB regelmatig te monitoren.

HBV genotype informatie is belangrijk en geassocieerd met response op de behandeling en voorspelling van het beloop van de ziekte. Genotype B en genotype A zijn geassocieerd met betere response op de behandeling in vergelijking met het genotype C en het genotype D. Vergelijkbare resultaten zijn gevonden bij "pegylated-interferon" behandeling voor deze genotypen; dus de bepaling van het HBV genotype is voornamelijk nodig om de uitkomst van de behandeling en de prognoses van de ziekte te kunnen voorspellen zoals aangegeven volgens de richtlijnen in de verschillende Europese landen inclusief Nederland, Zweden en Duitsland. Drie internationale organisaties: AASLD, EASL en APASL, hebben dringend verzocht om de rol van het HBV genotype over een groter gebied in de klinische praktijk vast te stellen. Het HBV genotype geeft belangrijke informatie in aanvulling op het serologische profiel en moleculaire karakteristieken.

HBV genotypering door middel van sequencing wordt als gouden standaard beschouwd. Echter, dit is niet de meest geschikte methode om op grote schaal genotypering te doen. In onze studie zijn single Nucleotide Polymorphisms (SNPs) ontdekt in een hoog variabel deel van preS2 en deze kunnen worden gebruikt als een eenvoudig genotyperingsmethode met behulp van RFLP (**Hoofdstuk 7**). Een simpel HBV genotyperingssysteem door middel van de RFLP methode is ontwikkeld voor het classificeren van de genotypen A tot H (**Hoofdstuk 7**). Dit nieuwe RFLP systeem is eenvoudig en gebruiksvriendelijk om snel en sensitief het HBV genotyping te detecteren in centra die geen sequencing faciliteiten hebben. De methode kan toegepast worden voor klinisch onderzoek met implicaties voor de behandeling van HBV patiënten en voor epidemiologische studies.

Conclusie

Het hepatitis B virus dat in Nederland voorkomt - vooral bij drugsverslaafden, bij degenen die met HIV geïnfecteerd zijn en bij allochtonen van niet westerse afkomst - is een virus dat via bloedcontact en van moeder op kind overdragen wordt. In het Verre Oosten komt het virus veel vaker voor en daar is 30 – 50% van de bevolking geïnfecteerd. Hepatitis B kan verschillende ziektebeelden veroorzaken: van een milde infectie tot een zeer ernstige levensbedreigende infectie, van een passagère infectie zonder klachten tot leverkanker. Het virus blijft vaak levenslang in het bloed aanwezig. In dit proefschrift is beschreven dat het virusgenoom zeer snel muteert; er zijn vele

genetische variaties van het Hepatitis B virus. Die variaties worden niet allen veroorzaakt door de mutaties maar ook omdat het virus een reparatie mechanisme ontbeert waardoor mutaties niet hersteld kunnen worden. Al die genetische verschillende vormen (genotypen) komen verspreidt bij de mens voor. Maar in dit proefschrift is aangetoond dat bevolkingsgroepen in Zuidoost Azië en de Pacific meestal een eigen genotype van het hepatitis B virus bij zich dragen. Zo is de verdeling van de HBV genetische varianten in Indonesië geassocieerd met de etnische achtergrond en de geografische herkomst van de gastheer. Ook de immunologische druk door de gastheer draagt bij aan de genetische diversiteit van HBV. (De patiënt maakt antilichamen tegen het virus en het virus tracht aan de antilichamen te ontsnappen door mutaties). De genetische diversiteit wordt dus veroorzaakt door virus – gastheer interacties en loopt parallel aan de evolutie van de mens in geografische verschillende gebieden. Sommigen van deze HBV varianten hebben een grote invloed op de volksgezondheid. Mutaties in het Hepatitis B virus (vooral in het z.g. HBsAg) kan leiden tot het ontstaan van virusvarianten die niet gedetecteerd worden in de commercieel beschikbare analyses. Hierdoor kan in donorbloed dat routinematig op de aanwezigheid van hepatitis B virus wordt gescreend, toch hepatitis B varianten voorkomen, die aan de screening zijn ontsnapt. Dit kan zeer ernstige gevolgen hebben voor de volksgezondheid. Deze alarmerende data zijn aanleiding om het nationaal beleid betreffende bloeddonaties te herzien. Artsen zijn voorzichtiger geworden met het geven van bloed of bloedproducten.

Ten slotte, in dit proefschrift is aangetoond dat de genetische variabiliteit van het hepatitis B virus parallel loopt met de etniciteit en geografische woonplaats van een gemeenschap en met dialecten en talen die in de gemeenschap gesproken wordt. De kennis van deze varianten heeft geleerd dat bepaalde varianten virulenter zijn, zich gemakkelijker verspreiden en minder gemakkelijk zijn te bestrijden.

De HBV variatie kan gebruikt worden voor bestudering van etniciteit, sociologie, archeologie van een bepaalde bevolkingsgroep.

Tot slot, levert de bestudering van deze varianten een essentiële bijdrage aan de ontwikkeling van diagnostische bepalingen en vaccins.

Ringkasan

Ringkasan

Infeksi oleh virus hepatitis B (VHB) atau hepatitis B merupakan masalah kesehatan yang penting di seluruh dunia. Hepatitis B kronis mengakibatkan macam-macam komplikasi dari yang tanpa gejala (*asymptomatic*), sirosis hati, gagal hati, hingga karsinoma hepatoseluler. Sehingga, hepatitis B kronis menjadi ancaman bagi kesehatan masyarakat dunia karena menyebabkan morbiditas dan mortalitas yang signifikan dan kerugian ekonomi yang cukup besar. Hal ini terutama menjadi masalah serius di regio Asia-Pasifik yang diidentifikasi sebagai tempat tinggal bagi kurang lebih 75% dari jumlah total penderita hepatitis B kronis. Beberapa faktor telah berhasil diidentifikasi sebagai hal yang berperan dalam manifestasi dan perjalanan penyakit hepatitis B kronis, termasuk variasi genetik VHB dan pejamu (inang). Sehingga pemahaman dan pengetahuan tentang variasi genetik VHB menjadi hal penting dalam mengembangkan strategi pencegahan dan pengobatan infeksi VHB.

VHB merupakan virus DNA untai ganda dengan 4 *Open Reading Frames* (ORFs), yakni surface (S), core (C), polymerase (P), dan x gen, yang keempatnya saling tumpang tindih. Kondisi genom yang kompak ini dengan laju mutasi tinggi dan tidak mempunyai kemampuan *proof reading,* dikombinasi dengan pengaruh imun sistem pejamu mengakibatkan variasi genetik yang tinggi pada genom VHB yang berkontribusi pada masalah kesehatan masayarakat dan klinikal. Hal tersebut menekankan pentingnya pengetahuan dan pemahaman karakteristik genom VHB dan patogenesis molekuler dalam menurunkan morbiditas dan mortalitas.

Variasi genetik VHB mengakibatkan VHB dapat diklasifikasikan menjadi 8 genotip, genotip A hingga H, dengan 2 genotip baru yang ditemukan di Asia yakni genotip I dan J. Menariknya, distribusi genotip VHB bersifat geografik dan etnik khususnya di Indonesia seperti yang dibahas dalam **Chapter 2** dan juga di tingkat regional Asia-Pasifik seperti yang dibahas dalam **Chapter 3**. Pada Chapter 2 disebutkan bahwa genom VHB menunjukkan karakteristik yang berbeda signifikan di seluruh Indonesia dengan subgenotip B3 dominan terutama di kepulauan bagian barat, sedangkan 5 subgenotip B lainnya (B3, B5, B7, B8, dan B9) dan 2 subgenotip C (C1 dan C2) dominan terutama di kepulauan bagian timur, dan C6 yang hanya ada di Papua, yakni pulau yang terletak paling timur di Indonesia. Data menarik lainnya yang ditemukan adalah observasi bahwa B2, yang spesifik pada populasi Cina di Asia daratan dan Taiwan, juga dominan pada populasi Cina di Indonesia.

Fenomena unik ini juga didukung oleh studi lain yang mempelajari genetik karakteristik VHB genotip C yang ternyata juga dipertahankan pada populasi tertentu seperti yang

dibahas dalam **Chapter 3**. Isolat VHB pada populasi dengan bahasa Austronesia pada Asia Timur hingga Asia Tenggara menunjukkan karakteristik berbeda dibandingkan dengan isolat pada populasi Papua yang berbahasa Papua di Papua, Papua New Guinea (PNG), hingga Pasifik. Perbedaan ini tidak hanya terletak pada pola genetik VHB tetapi juga pada karakteristik *immunoepitope* VHB. Sehingga, berdasarkan karakteristik molekuler, *immunoepitope*, geografi, dan asal etnik pejamu, VHB genotip C dapat diklasifikasikan menjadi 2 tipe, tipe Asia dan Papua-Pasifik. Hal penting lain yang didapat dari studi ini adalah distribusi karakteristik genetik VHB, terutama genotip B dan C, tampaknya membentuk suatu pola spesifik dari Asia Timur hingga Asia Tenggara di barat hingga ke Pasifik di timur. Pola ini membentuk *west* (subtype *adrq+*) *to east* (subtype *adrq-*) *gradient*. Uniknya, perubahan gradien ini diperkirakan terjadi di Papua dan PNG, dengan bukti ditemukannya bentuk peralihan subtipe *adrq*, yakni *adrq-indeterminate* A159/A177 and V159/V177. Sehingga, berdasarkan data subtipe, kami memperkirakan bahwa Papua dan PNG adalah daerah tempat terjadinya evolusi genom VHB genotip C dari tipe Asia ke tipe Papua-Pasifik.

Selain kontribusi pada epidemiologi, variasi genetik VHB telah pernah dilaporkan memberikan masalah kesehatan masyarakat dan masalah klinis. Adanya *occult hepatitis B* pada donor darah yang disebabkan karena mutasi pada gen *surface* pernah dilaporkan (**Chapter 4**), dan selanjutnya mengakibatkan mutasi pada protein HBsAg yang mengakibatkan penurunan antigenisitas dan konformasi protein *surface* (**Chapter 5**), sehingga HBsAg tidak dapat terdeteksi walaupun DNA VHB positif. Penemuan ini menekankan resiko mutasi HBsAg pada kesehatan masyarakat dan pada klinikal, sehingga masalah *occult hepatitis B* harus dipelajari lebih lanjut dan dikarakterisasi, dan strategi untuk melakukan skrining serta manajemen kondisi individu dengan HBsAg negatif dan anti-HBc positif harus terus dikembangkan.

Mutasi pada precore dan basal core promoter (BCP) telah pernah dilaporkan dan berkaitan dengan progresifitas penyakit dan insidens karsinoma hati seluler yang meningkat. Mutasi precore perubahan G ke A pada nt 1896 (G1896A) membentuk stop kodon pada kodon 28, sedangkan mutasi kritikal pada BCP (nukleotida 1742-1849) timbul dengan perubahan A ke T pada nt 1764 (A1762T/G1764A). Pada studi yang dilakukan pada pasien hepatitis B kronis (**Chapter 6**), kami menemukan bahwa prevalensi mutasi precore A1896 lebih banyak ditemukan pada usia lebih tua dan stadium penyakit lebih lanjut. Sementara itu, mutasi BCP T1762/A1764 ditemukan dengan jumlah yang kurang lebih sama pada semua fase hepatitis B kronis. Penemuan ini mengindikasikan bahwa mutasi ganda ini bisa timbul pada setiap waktu pada pasien hepatitis B kronis. Mutasi BCP ditemukan lebih sering pada genotip C/subtipe *adr* dibandingkan genotip B/subtype *adw*. Selain itu, studi ini juga menemukan bahwa kadar

HBsAg sama pada semua fase hepatitis B kronis, dan kadar HBsAg tidak berkorelasi dengan kadar DNA VHB. Hal ini mencerminkan bahwa aktifitas transkripsional cccDNA VHB timbul pada seluruh fase, dan tidak berkorelasi dengan kadar DNA VHB. Data di atas menunjukkan pentingnya untuk melakukan *monitoring* pasien-pasien dengan hepatitis B kronis.

Informasi genotip VHB sangatlah penting karena berkaitan dengan respons pengobatan dan prognosis penyakit. Genotip B dan A berkaitan dengan respons yang lebih baik terhadap pengobatan nukleosida dibandingkan genotip C dan D. Hasil yang sama ditemukan pula pada pengobatan dengan *pegylated-interferon*; sehingga informasi genotip VHB sangatlah penting untuk prediksi respons pengobatan dan prognosis penyakit, seperti yang juga direkomendasikan oleh beberapa *guidelines* dari negara-negara Eropa termasuk Belanda, Swedia, dan Jerman. Tiga badan regional khusus penyakit hati, AASLD, EASL, dan APASL telah merekomendasikan untuk melakukan pemeriksaan genotip VHB bagi keperluan klinikal. Sehingga, pemeriksaan genotip VHB sangatlah penting disamping pemeriksaan profil serologi dan karakteristik molekuler.

Hingga kini, pemeriksaan genotip VHB menggunakan analisa pohon filogenetik merupakan pemeriksaan *'gold standard',* namun karena pemeriksaan ini mahal dan lama, maka bila dilakukan pada jumlah sampel yang banyak akan membuat biaya menjadi sangat tinggi dan waktu lama sehingga kurang cocok untuk pemeriksaan dalam skala besar. Studi kami pada **Chapter 7** menunjukkan bahwa variasi nukleotida atau *Single Nucleotide Polymorphism* (SNPs) daerah preS2 genom VHB dapat digunakan untuk pemeriksaan genotip VHB dengan metode *Restriction Fragment Length Polymorphism* (RFLP) menggunakan ensim restriksi tertentu. Pemeriksaan ini merupakan pemeriksaan yang sederhana, sensitif, mudah, dan murah yang dapat mengklasifikasikan VHB menjadi 8 genotip. Keuntungan lainnya adalah pemeriksaan ini mudah digunakan pada penelitian dengan jumlah sampel yang banyak dan terbatasnya fasilitas peralatan penelitian.

Kesimpulan

Distribusi keaneka-ragaman genetik VHB di Indonesia berkaitan dengan etnisitas dan geografi pejamu. Faktor evolusi internal genom VHB termasuk rekombinasi antar *strain* dan tekanan imun sistem pejamu merupakan faktor yang diprediksi sangat berperan dalam keaneka-ragaman ini. Sehingga terdapat dua hal penting, yakni interaksi virus dan pejamu, dan evolusi paralel virus-pejamu pada geografi yang berbeda. Beberapa

keaneka-ragaman ini telah diidentifikasi mengakibatkan masalah kesehatan masyarakat dan masalah klinikal. Misalnya mutasi pada protein HBsAg terutama pada *determinant 'a'* mengakibatkan HBsAg tidak dapat terdeteksi oleh kit diagnostik komersial yang ada. Mutasi ini dapat menyebabkan *occult hepatitis B,* tidak terdeteksinya HBsAg sehingga dapat terjadi kesalahan manajemen penyakit. Hal ini terutama penting pada skrining donor darah, sehingga dapat terjadi kemungkinan lolosnya darah dan atau produk darah ke resipien. Sehingga sudah seharusnya kebijakan nasional untuk skrining donor darah yang berdasarkan HBsAg dikaji ulang bagi kepentingan masyarakat. Lebih lanjut, implikasi ini juga menjadi informasi yang berharga bagi klinikus agar berhati-hati dalam memberikan transfusi darah dan atau produk darah bagi pasien.

Pengetahuan keaneka-ragaman genetik VHB dan pengetahuan sistem imun pejamu sangatlah penting untuk mengerti patogenesis infeksi VHB dan dapat dilakukan pencegahan dan manajemen pasien lebih baik sehingga menurunkan angka morbiditas dan mortalitas. Selain itu, data keaneka-ragaman genetik VHB ini dapat dijadikan data dasar bagi pengembangan vaksin yang *ethnic-specific* dan sebagai informasi penting bagi pengembangan kit diagnostik. Secara paralel, pengetahuan variasi genetik VHB dapat diintegrasikan dan dihubungkan dengan perspektif linguistik, *human genetic*, dan arkeologi, yang diharapkan dapat meningkatkan derajat kesehatan manusia.

Acknowledgements – Dankwoord

The study was started when Prof Sangkot Marzuki motivated me about the importance of completing the formal degree and become a doctor a few years ago. This was what encouraged me to take the opportunity, and be a PhD candidate in University Medical Center (UMC) Utrecht, Netherlands, supervised directly by Prof David Handojo Muljono in Hepatitis Unit of Eijkman Institute, Jakarta. Therefore, this study was a collaboration study between the Eijkman Institute for Molecular Biology, Jakarta-Indonesia, and the Eijkman-Winkler Institute UMC, Utrecht-Netherlands. In this very special event, I would like to express my highest gratitude and my respect to Prof Sangkot Marzuki, who offered me this opportunity and was very supportive in the whole study. I am very grateful to Prof David Handojo Muljono for his direct supervision and continuous support from the beginning till the whole study is completely done. I am very much indebted to Prof Jan Verhoef for the opportunity, support, and collaborative work done in Eijkman Winkler Institute, Utrecht, and for our discussions. I am very grateful to Prof Andy Hoepelman for the invaluable discussions and supports.

I am really grateful to Greet Boland for her direct supervision for the part of the study done in the Eijkman Winkler Institute, Utrecht, and for our nice friendship. I am grateful to Harm Snippe for his support especially when part of this study done in Utrecht. My grateful goes to Prof Herawati Sudoyo for her support that allowed me to take this opportunity. I am very gratitude to Yfke Joseph for her assistance from the beginning of the collaboration until the study was completely done, and for her hospitality and for our friendship. I deeply thank Machiel de Vos who kindly helped me to find a lot of samples in Eijkman Winkler Institute, Utrecht. My deep thanks goes also to Dr Kwok for giving me blood samples from Suriname.

I do really thank to all my colleagues in Hepatitis Unit, Turyadi, Caecilia Hapsari, Korri, Susan, Erick, Santy, and dr.Martono. Thanks for all your support to completing this study. Personally thanks to Korri who will take part as a paranimf. My deeply thank to Rahmi Safitri for her assistance to arrange everything related to our activities in Hepatitis Unit and accomplish the printing of the thesis book. I deeply thank Jan and Henny van Kast for their continuous support and encouragement to completing the study, and especially for Henny for her Indonesian's food. To Theo and Will, I deeply thank for your support especially when I was in Utrecht. My special thank goes to Rowena for her support when the part of work done in Utrecht, and for our friendship. My thanks also to Risma Ikawaty, Mba Ina, Mba Tari, and Mba Helena for your support and chats. My special thanks to Dr Benny Phillippi for his support my research and my formal education. I send my gratitude to the Board of St Carolus Foundation and to the Board of Director of St Carolus Health Services for all the support that allowed me to get this opportunity.

It is my great pleasure to thank my family for all the endless support and prayer, and to my beloved parents, my father and my mother in heaven, thank you so much for teaching me to be who I am, and for your endless support and prayer. I send my thanks to all my brothers and sisters, who endlessly encourage me with their prayer and support to my family when I was away from home for months. My apology if I can not mention a lot of friends or colleagues in this page who contributed indirectly to this thesis. And, my highest gratitude goes to the God Almighty Jesus Christ for giving me this opportunity, for His bless and love.

CURRICULUM VITAE

The author of this thesis was born on 24 April 1962 in Jakarta, Indonesia. In 1990, she finished her study in medicine from Medical Faculty of Tarumanagara University in Jakarta. Thereafter, she spent three years as a medical doctor in a hospital in West Jakarta, and then she took apart as a medical doctor for three years in a district general hospital in Tarakan island, East Kalimantan. In 2000, she joined the Eijkman Institute, Jakarta, as a master student where she started her research in Hepatitis Unit under supervision of Prof dr. David H. Muljono, SpPD, PhD. In 2002, she finished her master degree in Immunology of Biomedical Science of Indonesia University. Her interest in molecular biology particularly in hepatitis research has been led her to join the Eijkman Institute after completing her master degree. To perform her medical practices, since 2002 she has been working as a medical doctor in St Carolus Health Services, Jakarta.

List of Publication

1. **Thedja** MD, Muljono DH, Nurainy N, Sukowati CH, Verhoef J, Marzuki S. Ethnogeographical structure of hepatitis B virus genotype distribution in Indonesia and discovery of a new subgenotype, B9. Arch Virol. 2011 May;156(5):855-68.

2. **Thedja** MD, Roni M, Harahap AR, Siregar NC, Ie SI, Muljono DH. Occult hepatitis B in blood donors in Indonesia: altered antigenicity of the hepatitis B virus surface protein. Hepatol Int 2010;4:608 – 614.

3. Ie SI, **Thedja** MD, Roni M, Muljono DH. Prediction of conformational changes by single mutation in the hepatitis B virus surface antigen (HBsAg) identified in HBsAg-negative blood donors. Virol J 2010;7:326.

ISBN 9789039358542