Host genetic factors in hepatitis B infection, liver cancer and vaccination response:
A review with a focus on Africa

Branwen J. Hennig*, Andrew J. Hall

Department of Epidemiology and Population Health, London School of Hygiene & Tropical Medicine, Keppel Street, London WC1E 7HT, UK

* Corresponding author
   Branwen J. Hennig
   MRC International Nutrition Group
   EPH
   London School of Hygiene & Tropical Medicine
   Keppel Street,
   London WC1E 7HT
   UK
   Email: branwen.hennig@lshtm.ac.uk
   Tel: +44 (0)20 79588131
Abstract
The disease burden due to hepatitis B virus (HBV) infection remains significant; 350 million people are infected world-wide, and around half a million deaths each year are due to HBV-related liver disease and hepatocellular carcinoma (HCC). Infant immunisation against infection was introduced in the early 1980s, the vaccine is routinely administered across regions where the disease is endemic and has been shown to be safe and effective. However, the large number of older individuals with persistent infection means that disease will not be reduced significantly for several decades. Furthermore, failure to respond to the vaccination has been observed in about 5% of vaccinees and to date we have limited information on the durability of vaccine protection against infection.
Hepatitis B infection and disease pathogenesis are known to be influenced by a number of factors including host genetics factors. This review aims to give an overview of the role of genetic variation in persistent HBV infection and the development of liver disease including HCC. Vaccine-induced immunity is, at least in part, heritable and we also discuss findings on the genetic control of responses to HBV vaccination.
The epidemiology of HBV infection differs by world region, as does the genetic makeup of individuals originating from different regions. This review focuses on the situation in Africa, where hepatitis B is highly endemic.

Key words: Host genetic variability, HBV, persistence, HCC, vaccine-induced immunity
Introduction

Hepatitis B virus infection and subsequent disease may take several decades to develop. Throughout the life course susceptibility to infection, disease progression and development of end stage liver disease may all be modified by factors such as the viral genotype, age at infection, mode of transmission, vaccine efficacy, reactivation / persistence, presence of other infections or co-factors (e.g. aflatoxin, alcohol) and immune competence/suppression (Hall et al., 2002). Any or all of these may be affected by host genetic factors but the role of these is still poorly understood. Yet, taking a genetic approach (in conjunction with clinical and epidemiological methods) may have profound implications for our understanding of disease pathogenesis, molecular diagnosis, identification of novel treatments, and gene-environment interactions, all of which may lead to better preventive strategies. In addition controlling for genetics may allow the recognition of previously unknown environmental factors leading to improved prevention. This review aims to summarise our current understanding of the role of host genetic factors in persistence of HBV infection, disease progression and vaccine-induced immunity.

Hepatitis B is highly endemic across many parts of Africa, Asia and in the Amazon basin of South America; however the epidemiology of hepatitis B virus (HBV) infection differs by region, as does the genetic makeup of individuals originating from these regions. This review concentrates mainly on the African context. Estimates by the World Health Organisation (WHO, 2008) and the Global Alliance for Vaccines and Immunization (GAVI, 2005) show that around two billion people, i.e. one-third of the global population, have been infected with HBV at some stage in their lifetime. In Africa transmission occurs in early life, mainly from child-to-child probably by a variety of routes, relatively rarely from mother to child and in a second wave later in life through sexual contact. Infection at birth leads to persistent (lifelong) infection in about 90% of individuals, in childhood to about 25-40%, but infection as an adult results in less than 10% persistent infection. This infection is defined by the presence of hepatitis B surface antigen (HBsAg) in blood and those affected have an up to 25% risk of dying prematurely from cirrhosis (scarring of the liver) and/or hepatocellular carcinoma (HCC), although the risk in females is substantially lower than in males [pic](Bah et al., 2001; Dickinson et al., 2002; Yu et al., 2000a).

It is thought that HBV or viral proteins may play a role in HCC induction, that prolonged inflammation during infection accelerates hepatocarcinogenesis and that viral integration into the hepatocyte genome is important (Hino, 2005).

The prevalence of persistent HBsAg carriage, as shown in Figure 1 (men only), and consequently HCC varies widely even within the African continent and implies higher transmission in childhood in dry zones (Hall, 2004). In Africa HBV infection rates equate to about 60 million persistently infected people. Liver cancer - which is largely attributable to HBV – is among the most common cancers (11.5 and 5.0% in males and females, respectively), and primary liver cancer accounts for an estimated 200,000 deaths in Sub-Saharan Africa every year [pic](Parkin, 2006; Parkin et al., 2008’). However, it has to be emphasised that we are still lacking good prevalence and incidence data on many chronic diseases in Africa, including HBV infection and chronic liver disease.

It is clear that Hepatitis B remains one of the major diseases of mankind and poses a serious global health burden, especially in Sub-Saharan Africa. Treatment of HBV infection with interferon-(, lamivudine, and more recent anti-virals is possible - but costly and thus generally not available in Africa.

However, HBV infection is preventable through vaccination, which has been available since 1982,
and immunization was recommended by the World Health Assembly in 1992 to be introduced in all countries. Vaccination coverage across the African continent is variable, as shown in Figure 2. Overall, large-scale vaccination has been demonstrated to be safe and cost-effective given the high level of vaccine efficacy. Vaccine-induced immunity is affected by the type and administration route of the vaccine itself, age, gender, UV light exposure, smoking, co-infections, and nutritional factors (reviewed by [24](Van Loveren et al., 2001). The peak anti-HBs level attained after vaccination can be used as a surrogate for vaccine efficacy (i.e. protection against infection and persistent carriage) long-term. Antibody decays in an exponential manner over time irrespective of the population under study (Jack et al., 1999; van der Sande et al., 2006) although protection appears to remain in many who have lost antibody. Non-response to vaccination is observed in approximately 5% of vaccinees (Zuckerman, 2006).

This review describes genetic analysis approaches to hepatitis B virus persistence, related liver cancer and vaccination response with a focus on studies in African populations (based on publications in English or French). Family linkage and association analysis aim to map the chromosomal location(s) of the disease causing variant(s) in families with multiple affected individuals. Twin studies provide a means of distinguishing the risk attributable to genetic and environmental factors. Population based (case-control) association studies are designed to determine if there is a correlation between an allelic variant and any given outcome, e.g. disease status. Hepatitis B has been shown to cluster in families, thus it has been possible to collect evidence for a role of genetic factors in HBV infection, progression and vaccine responses from family, twin and population based approaches.

Definitions
HBV infection is defined by the presence of anti-core antibodies (anti-HBc). IgM anti-HBc reflects recent infection and IgG anti-HBc prior or active infection. Individuals with persistent infection are in addition surface antigen (HBsAg) positive. HBV DNA is usually detectable in persistent carriers and indicates viral replication in liver cells. The presence of circulating e antigen (HBeAg) correlates with HBV DNA levels and therefore also indicates viral replication with associated infectivity. Antibodies to surface antigen (anti-HBs) appear on recovery from infection or are mounted in response to HBV vaccination. For further details see Decker (Decker, 1998). Positivity for anti-core antibody in the absence of other serological markers, “anti-HBc alone” is compatible with either acute resolved infection or with persistent HBV infection (Grob et al., 2000).

Genetic susceptibility to persistent HBV infection
Family clustering of persistent HBV carriage has been observed in many geographical settings [25](Beasley et al., 1983; Blumberg et al., 1969; Bosch et al., 1973; Lin et al., 2006; Motta-Castro et al., 2005; Obayashi et al., 1972; Stroffolini et al., 1991). However, there appears to be only one report of a family-based linkage study from The Gambia [26](Frodsham et al., 2006). This study was based on the analysis of a panel of over 300 microsatellite markers in 162 independent affected sibling pairs in 135 families and identified a region of linkage on chromosome 21q22 with a logarithm of odds (LOD) score of 3.16 (P<0.001). This region contains the cluster of cytokine class II receptor (CRF2) genes, including IFNAR1, IFNAR2, IL10RB and IFNGR2. The analysis of a further 22 mostly single nucleotide polymorphisms (SNPs) revealed associations with persistent HBV infection (HBsAg + / anti-HBc-IgM -) of two non-synonymous
coding changes in IFNAR2 (F10V; rs1051393) and IL10RB (K47E; rs2834167), respectively, and in vitro assays supported a functional effect of these variants in receptor signaling.

We found only one twin study regarding susceptibility to HBV infection based on a Chinese population (Lin et al., 1989). Lin and co-investigators demonstrated differences in concordance of infection in monozygotic (MZ) twins compared to controls and dizygotic (DZ) twins and controls (but no difference between MZ and DZ) as well as differences in carrier status between MZ and DZ and MZ and controls (but not DZ and controls), suggesting the importance of maternal infection in this population, and that there may be some effect of host genetic factors on the progression to persistent carrier status.

The bulk of work on genetic susceptibility to HBV infection comes from case-control association studies, with most of the earlier studies concentrating on variation in the major histocompatibility complex (MHC) located on chromosome 6. MHC class II molecules are expressed on antigen presenting cells and involved in presentation of pathogens such as HBV virus particles to T helper cells, thus triggering adaptive immune responses. In Gambian children and adult males the DRB1*1302 allele was shown to be associated with spontaneous elimination of infection (HBsAg - / IgG anti-HBc +) (Thursz et al., 1995), and this finding has since been replicated in several populations of Caucasian origin (Hohler et al., 1997; Thio et al., 2003). However, this was not confirmed in a more recent study population from Togo and Benin, although sample numbers in this subgroup analysis were very low (Bronowicki et al., 2008). In the same study DRB1*08 was correlated with persistent infection in “HBc alone” subjects and DRB1*09 appeared associated with the presence of anti-HBs antibodies, a result that stands in contrast to the findings from a Korean study (Ahn et al., 2000). DQA1*0501 and DQB1*0301, but not other human leukocyte antigen (HLA)-A, -B, or -DRB1 alleles, were associated with persistence (HBsAg + / anti-HBc +) in African-Americans from the East coast of the USA (Thio et al., 1999).

The HLA-A, -B, and C antigen distribution was also assessed in HBsAg positive/negative and HBeAg positive/negative Senegalese individuals by Dieye and colleagues (Dieye et al., 1999; Obami-Itou et al., 2000). Associations between A1, A23, B8 and Cw3 and HBsAg positivity as well as A1 and HBeAg positivity were found, suggesting an influence on persistence as well as infectivity. The findings relating to variation in the MHC region suggest that further studies are necessary to establish which specific alleles affect susceptibility to HBV persistence.

HLA studies based in non-African populations (e.g. Europe, Qatar, China, Korea, USA) have been reviewed extensively (de Andrade and de Andrade, 2004; Frodsham, 2005; He et al., 2006; Thursz, 2001a; Thursz, 2001b; Wang, 2003). To date one genome-wide association study (GWAS) has been published, reporting the HLA-DPA1 and HLA-DPB1 locus to be robustly associated with persistent HBV infection in Japanese and Thai (Kamatani et al., 2009). The authors suggest that antigen presentation on HLA-DP molecules might be critical for virus elimination. The finding from this GWAS emphasizes the role of variation in MHC genes in HBV disease progression.

Studies on non-MHC genes and susceptibility to and progression of HBV infection are limited, in particular with regard to those based in African populations or those of African origin; reports are available on: TNFA (tumour necrosis factor alpha), MBL2 (mannose binding lectin, also known as MBP), VDR (vitamin D receptor), CTLA4 (cytotoxic T-lymphocyte antigen 4), MTHFR (5,10-methylenetetrahydrofolate reductase), MTR (5-methyltetrahydrofolate-homocysteine methyltransferase) (Bronowicki et al., 2008) as well as IL10, IL19 and IL20 (Truelove et al., 2008). Thursz et al. described that the -308 promoter polymorphism (rs1800629) in TNFA was
associated with an approximately 2-fold increased risk of persistent HBV infection (HBsAg +, total anti-HBc +, IgM anti-HBc -) in The Gambia (Thursz et al., 1996). The functional relevance of this variant remains unclear, but it is thought to affect TNF- expression levels and may thus influence levels of this pro-inflammatory cytokine, which could have an impact especially at the early stages of infection. The same Gambian population (although a smaller subset) was also studied for variation in the MBL2 (Bellamy et al., 1998). This protein is important in innate immunity due to its role in opsonisation and phagocytosis and due to the presence of changes in the coding region of the gene that affect protein serum levels, however, no association was reported with HBV infection or persistence. In a further study based on Gambians, a polymorphism in the VDR gene was genotyped and found to correlate with persistent carriage (HBsAg + / anti–HBc +) (Bellamy et al., 1999). This is an intronic SNP (rs731236) and thus not likely to be of functional relevance, although the VDR has been shown to affect immune-regulatory processes and thus may be important in viral infection. Finally, the screening of six SNPs across the CTLA4 gene in African-Americans by Thio and colleagues revealed single SNP and haplotype associations with infection and carriage (Thio et al., 2004). In particular +49G/X (rs231775) was detected more often in persons who recovered from infection and -6230A (rs3087243) more frequent in individuals with viral persistence. CTLA4 is a protein which transmits inhibitory signals to T cells and it is thought to be crucial in immune regulatory processes. Variation in determinants of homocysteine metabolism, which may affect viral inactivation through DNA methylation processes and thus outcome of HBV infection, was investigated in a population from Togo and Benin by Bronowicki and colleagues who screened markers in the MTHFR and MTR genes (Bronowicki et al., 2008). SNPs MTHFR C677T (rs1801133) and A1298C (rs1801131), as well as MTR A2756G (rs1805087) frequencies were compared between HBsAg positives and anti-HBs positives (all anti-HBc +). The presence of MTHFR 677T or MTR 2756G and Beninese origin was found to predict anti-HBs seroconversion. The comparison between anti-HBs positive and all anti-HBs negative subjects (i.e. including anti-HBc negatives) only showed MTHFR 677T, but not MTR 2756G, and Beninese origin to be correlated with presence of antibodies to surface antigen. Finally, Truelove and colleagues assessed 25 SNPs in IL10, 10 in IL19 and 7 in IL20 in a case-control cohort of 398 African Americans (and European Americans) with persistent or self-limiting HBV infection, matched by HIV status (Truelove et al., 2008). Two markers in the IL10 promoter (including rs1800896 at position -1082) and an intergenic SNP near IL20 were associated with persistent HBV infection, whilst an intronic IL10 SNP and another intergenic SNP near IL20 correlated with viral clearance; an IL20 haplotype was also associated with susceptibility to persistent infection (for details see table 1). IL19 and IL20 both belong to the IL10 family of genes (which are known immune regulatory molecules), and are thought to share receptor chains between each other and with IL10, thus potentially also sharing signalling pathways. The results presented by Truelove et al do not point towards a functional marker in IL10 or IL20 and further fine-mapping would be needed to identify any causative variant. It seems that none of these reports on non-MHC gene variation has been followed-up in other populations of African origin to date.

Reports in populations not of African origin have been published mostly on genes that modulate or control the immune response to HBV infection. Several reviews have summarized these (de Andrade and de Andrade, 2004; Frodsham, 2005; He et al., 2006; Sun et al., 2009; Wang, 2003), and include: TNF, CCR5, CTLA4, RANTES, MCP1, ESR1, IL10, IL18, IL19, IL20, KIR, FAS and its ligand, TBX21, CD14, CCR5, VDR, MBL2, IFNG and receptors, IFNA and receptors etc. A recent meta-analysis found no association between HBV infection status and the
TNFA rs1800629 SNP at position -308 (Qin et al., 2010) in Asians and Europeans; yet an earlier meta-analysis did correlate this variant with persistent HBV infection in Asians, but not Europeans (Zheng et al., 2010).

**Genetic factors in HBV-related liver disease and HCC**

Cases of HCC cluster in families (Yu et al., 2000b), probably reflecting the clustering of persistent carriers (see above), given that HCC is etiologically associated with HBV in 80% of cases in high endemicity areas and that the majority of deaths in HBV carriers is due to HCC [Dickinson et al., 2002; Parkin, 2006; Yu et al., 2000a]. There are a number of studies that have assessed the host genetic component in the development of HBV-related HCC, most of these relate to Asian populations.

As far as we are aware there is no genome-wide family based linkage study for HBV-related HCC, however, there is a publication which presents findings from linkage and family-based association analyses for chromosome 4 in Chinese families from Taiwan, reporting suggestive evidence of linkage to 4q22.3-28.1 [Shih et al., 2006]. Similarly, only anecdotal evidence is available regarding twins with Demir et al. describing an identical Turkish twin brother pair, diagnosed at the same time with HCC and both confirmed to be persistent HBV carriers (Demir et al., 2002).

Candidate genes for studies on liver disease relate to genes that affect immune responses to HBV infection; genes that metabolise hormones or xenobiotics, genes that modulate the development of fibrosis/cirrhosis, and genes thought to affect cancer molecular pathogenesis such as those controlling cell cycle control mechanisms, DNA repair, cellular motility, cancer predisposition [Furberg and Ambrosone, 2001; Kim and Lee, 2005]. There are also studies that have examined genotypes of enzymes thought to activate aflatoxin to the reactive epoxide form. This is important as there is a strong interaction between aflatoxin and hepatitis B in carcinogenesis [Sylla et al., 1999; Uwaifo and Bababunmi, 1984]. In the African context we identified several publications on genetics of HBV-related liver disease relating to Gambians, Ghanaians, Sudanese and Moroccans, respectively. Polymorphisms in enzymes involved in carcinogen metabolism (glutathione S-transferases encoded by GSTM1, GSTT1, and epoxide hydrolase 1 encode by EPHX1 (also known as HYL1 or mEH) and DNA repair (X-ray repair cross complementing protein encoded by XRCC1) were investigated in a hospital-based case-control study of HCC in The Gambia [Kirk et al., 2005]. Carriers of a GSTM1 deletion and those heterozygous for a coding change in the XRCC1 gene (Arg399Gln, rs25487) where more likely to suffer from HCC, and this effect was even more prominent in individuals exposed to aflatoxin. The authors concluded that genetic modulation of carcinogen metabolism and DNA repair appears to alter susceptibility to HCC and that these effects may be modified by environmental factors, i.e. aflatoxin exposure. An earlier study in rural Gambians had shown the GSTM1 null genotype to lead to increased aflatoxin-adduct levels in HBsAg negative, but not HBsAg positive individuals with HCC [Wild et al., 2000]. Dash and colleagues investigated the role of host genetic variation in GSTM1, GSTT1 and EPHX1 with regard to aflatoxin exposure (measured as aflatoxin albumin adduct level in blood) in Ghana (Dash et al., 2007). Employing multivariable analysis (adjusting for gender, genotypes, HBsAg status, and age) they found female gender and EPHX1 R139H (rs2234922) heterozygote status to associate with increased mean aflatoxin albumin adduct levels. This indicates that these individuals may be most susceptible to aflatoxin / HBV-related liver disease, although HCC is known to be less common in women [Bah et al.,
2001; Yu et al., 2000a). However, HCC was not a known outcome in this study population and a future follow-up study would be necessary to shed more light on these relationships between risk genotypes, gender, aflatoxin exposure and HCC. A group in Sudan also reported on correlations between peanut butter intake (i.e. aflatoxin exposure) and HCC patients, in individuals with the GSTM1 null genotype, however neither GSTT1 nor EPHX appeared to modify this association (Omer et al., 2001; Tiemersma et al., 2001). Further studies will be necessary to disentangle the relationship between aflatoxin exposure, host genetic factors and HCC by looking at two-and three-way relationships between these factors. Finally, manganese superoxide dismutase, encoded by the SOD2 gene, is involved in the destruction of radicals which are normally produced within cells and which are toxic to biological systems. A non-synonymous coding change (Val to Ala, rs1799725) has been shown to have functional effects on these processes and was recently studied in HCC patients in Morocco (Ezzikouri et al., 2008). HCC cases were more likely to be homozygous for the Ala allele compared to controls, and this association was even stronger in those with confirmed hepatitis C infection. HBV infection was also assessed and the proportion of HBsAg positives was about four times higher in the HCC group, however, the total number of persistent carriers was only 19 and thus too small for a subgroup analysis. It would be interesting to assess this SOD2 variant in a larger HBV-related HCC case-control study.

Host genetic studies on HBV-related liver disease including cirrhosis and HCC have been published on extensively in Asian populations and to a lesser extent in those of European origin, covering a large number of candidate genes including (and this is not an exhaustive list): CYP1A1, ERS1, FAS, GSTM1, GSTT1, IL10, IL1B, IL1RN, NAT2, TGFB, TFNA, XRCC1, hMLH1, XPD (reviewed by (Chen et al., 2005; Kim and Lee, 2005; Kirk et al., 2006; Sun et al., 2009)). Recently, a GWAS has been conducted in Chinese, identifying a region on chromosome 1p33.26 (KIF1B, UBE4B and PGD) as susceptibility locus for HCC in individuals persistently infected with HBV (Zhang et al., 2010). The meta-analysis by Qin et al on TNFA polymorphisms (see above) concluded that the -308 promoter variant is associated with liver cancer in Asians and Caucasians when patients were compared to healthy controls, but not when compared to HBV infection cases (Qin et al., 2010). This shows that although there is a large body of evidence of host genetic factors in HBV-related liver disease derived especially from Asian studies, there is a distinct lack of equivalent studies from Africa, despite the high prevalence of disease in both these regions.

Genetics of immunity induced by HBV vaccination

Family and twin studies indicate that there is a significant heritable component to HBV vaccine-induced immunity both within Africa and across the rest of the world (reviewed by (Kimman et al., 2007). Newport et al. demonstrated higher concordance in MZ compared to DZ twins in The Gambia, with heritability estimates ranging between 63-85% (Newport et al., 2004). The authors further estimated the contribution of HLA DRB1 variation to account for 15% (0-63%) of the heritability and thus suggested that genes outside this locus significantly influenced vaccine-induced antibody levels. No other family-based or twin study appears to have been carried out on samples from the African continent. Elsewhere, a twin study was conducted in German vaccinees and identified a haplotype based on variation in the promoter of the IL10 gene and HLA DRB1 alleles to correlate with response to HBV vaccination (Hohler et al., 2005). A follow-up of this study suggested that differences in the T cell recognition of peptide/MHC complexes are critical in T cell responsiveness to HBsAg (Kruger et al., 2005). A family based association approach
revealed that carriage of the HLA class III C4AQ0 allele lead to non- or slow response to HBV vaccination in Italian subjects (De Silvestri et al., 2001). Work by investigators in the USA showed that with respect to non-response HLA identical siblings were concordant whereas the majority of haplo- or non-identical siblings were discordant, they also presented evidence for linkage with the MHC locus (LOD 6.3) in the families screened (Kruskall et al., 1992).

As with genetic susceptibility to HBV infection, most population based investigations on hepatitis B vaccine responses have concentrated on the role of variation within the MHC region. Such studies in various populations have been reviewed in detail [pic](Kimman et al., 2007; Milich and Leroux-Roels, 2003). Briefly, it appears that alleles associated with high/normal response are: DRB1*01, DRB1*I1, DRB1*15, DQB1*0501, DPB1*0401 and DRPB1*0402; whereas alleles correlated with non- or poor response are: DRB1*03, DRB1*07, DQB1*02, DPB1*1101, DRB1*14, and DRQB1*020. The only HLA population-based study in individuals of African origin (other than the above twin study in Gambians) is that by Wang and colleagues who assessed vaccine-response genetics in an ethnically mixed US population comprising a large proportion of African-Americans (Wang et al., 2004). Non-response to vaccination was attributed to variants HLA-Cw*03, DRB1*07 and DQB1*02, whereas response was associated with DRB1*15 and DQB1*06 (all results adjusted for ethnicity and other covariates). Of these associations with DRB1*07 and DQB1*02 as well as DRB1*15 have been replicated (see above), but another previously replicated association with DRB1*03 and non-/poor response was not seen in the US based study.

Only a handful of reports on non-MHC genes have been published with respect to HBV-vaccine induced immunity, these describe the screening of genes including Hp (Louagie et al., 1993), GNB3 (Lindemann et al., 2002); IL2, IL4, IL6, IL10, and IL12B (Wang et al., 2004) TNFA, IL1B, IL2, IL2RA, IL4, IL4RA, IL10, IL12B, IL12RB1, IL12RB2, and IL13 [pic](Yucesoy et al., 2009). None of these are based on populations of native African origin. Wang and colleagues (Wang et al., 2004) studied a population of predominantly black African Americans and reported on a haplotype consisting of three IL4 SNPs (-1098G/T, rs2243248; -590C/T, rs2243250; and -33C/T, rs2070874) associated with good response to HBV vaccination, whereas those carrying a 4bp deletion in the IL12B promoter were shown to be non-responsive. In this sample HIV infection was also correlated with non-response to vaccination. IL4 is important in the modulation of B cell function, in particular immunoglobulin switching, and IL12B associates with IL23A to form the IL-23 interleukin, an heterodimeric cytokine which functions in innate and adaptive immunity. A possible implication of variation in these genes on HBV vaccine-induced immunity is therefore plausible. The study by Yucesoy et al [pic](Yucesoy et al., 2009) describes associations of a TNFA and a IL12B SNP with variations in median anti-HBs antibody levels; details given on the ethnicity of their study population is limited (84% of 141 non-Hispanic Whites), but a very small number of African Americans may have been included, thus we mention it in this review. A larger study comprising over 700 SNPs across a total of 133 candidate genes reported on associations in relation to immune responses in over 600 HBV infant vaccinees from The Gambia [pic](Hennig et al., 2008; Ryckman et al., 2010). This showed that variation in IFNG, MAPK8, IL10RA, CD44, CD58, CDC42, IL19, IL1R1, and to a lesser extent ITGAL affect peak anti-HBs level and that odds of core-conversion despite vaccination was associated with variation in CD163. Although these are plausible candidate molecules, the associated SNPs in these genes were all in non-coding regions and are thus unlikely to be of functional relevance.
(although they may be in linkage disequilibrium with functional variants). A coding change (R719T, rs2230433) in the ITGAL gene was the exception; this appears to lead to a good immune response in those homozygote for the Threonine allele. ITGAL forms a subunit of lymphocyte function-associated antigen-1 (LFA1), together with ITGB2, and plays a central role in immune cell interaction by binding to ICAMs and regulating leukocyte-endothelial cell interaction, cytotoxic T-cell mediated killing, and antibody dependent killing by granulocytes and monocytes. The largest study in this field to date is that by Davila et al (Davila et al., 2010), and comprised a two-stage screen of just under 6000 SNPs in over 900 immune genes with regard to HBV vaccine failure following a two-dose immunisation schedule in 981 Indonesians. Non-response was shown to be associated with SNPs in BTN2, IL6ST, KLRF1, LY6H, MBL2, TGFB3, HLAB, HLA-DRA, HLA-DRQB1, FOXP1, LILRB4, TGFB2, TNFSF15, C5, CCL15, with the most convincing evidence for MHC genes and the latter six genes listed. Unfortunately, a direct comparison between these two last mentioned larger studies (with respect to sample size and number of variants genotyped) is not possible due to differences in population, number of vaccine doses, age at vaccination and most importantly the timing and measurement of the primary outcome measure of vaccine-induced anti-HBs level. INFG and CD44 may possibly represent a common thread across majority of genes shown to affect HBV vaccine-induced immunity identified in both the Gambian and Indonesian study based on literature searches (Ryckman et al., 2010).

Conclusions
This review highlights how little we know about host genetic factors in hepatitis B infection/persistence, HBV-related liver disease including HCC, and immunity induced by hepatitis B vaccination. Genetic variation may affect these outcomes separately or jointly, given that there are likely to be commonalities in the immune responses to natural infection and vaccination and that immune modulation leading to persistent infection may directly or indirectly affect the development of HBV-related liver disease. The majority of studies published to date have assessed variation in the MHC region and there is some evidence that MHC class II alleles are important, but even here further research is needed. Data on non-MHC genes is at best scattered. In addition, many genetic studies discussed in this review are hampered by small sample sizes, the limited number of markers/genes screened, differences in study design and analysis, and the fact that results have not been replicated in most instances. There is furthermore a discrepancy between the reasonably high and very small number of reports from Asia and Africa, respectively. As mentioned above, although hepatitis B is highly endemic in both, the epidemiology of hepatitis B is different in Asia and Africa, so these regions should be assessed separately and/or comparatively. Within Africa it is notable that the majority of studies carried out to date are based on Gambians. This is likely because there has been a long-standing investment in hepatitis B work in The Gambia including the Gambian Hepatitis Intervention Study and the presence of a cancer registry there (Bah et al., 2001; Viviani et al., 2008).

Future research should comprise a variety of populations (especially within the African continent), so as to account for genetic variation between populations, but also differences in covariates and environmental factors, particularly with respect to aflatoxin exposure. More well-designed large-scale studies with long(er) follow-up times are needed, although intermediate (non-invasive) phenotypes, e.g. fibrosis stage measured by ultrasound as early indicator of disease progression, could also be assessed. GWAS would be useful for the identification of novel genes, i.e. helping
to generate and evaluate hypotheses of pathways and mechanisms underlying susceptibility to and pathogenesis of HBV infection, as well as responses to vaccination. Antigen presentation and recognition, the magnitude or kinetics of virus- or vaccine-induced antibody response, lymphocyte proliferation, and long-term immune memory are all processes that are likely to be affected by genetic variants, many of which are yet to be identified. Replication/transferability of linkage and association signals should also be followed up, e.g. a systematic screening of the MHC region for instance would be warranted for HBV infection and vaccination-related phenotypes. This would ideally also be followed by work looking into the functional relevance of genetic variation especially with respect to the MHC region and TNF. We are hopeful that with the advance of cheaper and more rapid genotyping and analysis methods as well as more multi-disciplinary and collaborative approaches, there will be greater opportunities to study the host genetics related to hepatitis B in the broadest sense and thereby ultimately to improve diagnosis, prognosis, treatment and prevention strategies of HBV-related pathogenesis.

Acknowledgements
BJH is grateful to Renato Mariani-Costantini, Mario Di Gioacchino and Nasr Eldin Elwali for their support and the opportunity to present this work during the EIDC 2008 meeting in Khartoum, Sudan.

References
GAVI. Hepatitis B factsheet, 2005.
He YL, Zhao YR, Zhang SL, Lin SM. Host susceptibility to persistent hepatitis B virus infection. World J Gastroenterol 2006; 12: 4788-93.
Hohler T, Reuss E, Freitag CM, Schneider PM. A functional polymorphism in the IL-10 promoter influences the response after vaccination with HBsAg and hepatitis A. Hepatology 2005;


Obayashi A, Okochi K, Mayumi M. Familial clustering of asymptomatic carriers of Australia antigen and patients with chronic liver disease or primary liver cancer. Gastroenterology


Thursz M. MHC and the viral hepatitides. Qjm 2001b; 94: 287-91.


Zheng MH, Qiu LX, Xin YN, Pan HF, Shi KQ, Chen YP. Tumor necrosis factor-alpha-308A allele may have a protective effect for chronic hepatitis B virus infection in Mongoloid populations. Int J Infect Dis 2010; 14: e580-5.

**Glossary**

GAVI Global Alliance for Vaccines and Immunization
HBV hepatitis B virus
HCC hepatocellular carcinoma
HLA human leukocyte antigen
LOD logarithm of odds
MHC major histocompatibility complex
MZ / DZ monozygotic / dizygotic twin
SNP single nucleotide polymorphism
WHO   World Health Organisation
Table 1: Summary of reports on genetic factors in susceptibility to persistent HBV infection, liver disease and HBV vaccine-induced immunity (as direct or indirect outcome measure)

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Locus / Gene investigated (number of markers genotyped)</th>
<th>Study population (N total)</th>
<th>Main finding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Susceptibility to persistent HBV infection</td>
<td>Genome-wide scan; Chr21q22 including IL10RB, IFNAR1, IFNAR2, IFNGR2 (22 additional markers)</td>
<td>The Gambia, families (N=182 families with one or more affected siblings; 192 unaffected siblings)</td>
<td>IFNAR2 F10V (rs1051393) (rs2834167) associated with persistence</td>
</tr>
<tr>
<td></td>
<td>HLA-DRB1</td>
<td>The Gambia, case-control (N=403 children; 135 adult males)</td>
<td>DRB1*1302 correlated with persistence</td>
</tr>
<tr>
<td></td>
<td>HLA-DRB1; MTHFR, MTR</td>
<td>Togo (N=305) and Benin (N=150), case-control</td>
<td>DRB1*08 correlated with persistence in “HBC alone” subjects</td>
</tr>
<tr>
<td></td>
<td>HLA-A, -B, and -C</td>
<td>Senegal, case-control (N=194)</td>
<td>A1, A23, B8 and Cw3 associated with positivity; A1 associated with positivity</td>
</tr>
<tr>
<td></td>
<td>HLA-A, -B, -DQA1, -DQB1, -DRB1</td>
<td>USA, case-control (N=190 African-Americans)</td>
<td>DQA1<em>0501 and DQB1</em>0301 associated with persistence</td>
</tr>
<tr>
<td></td>
<td>IL10 (25 SNPs), IL19 (10 SNPs), IL20 (7 SNPs)</td>
<td>USA, case-control (N=398 African-Americans)</td>
<td>IL10 (rs1800893, rs1800895, rs1518108) as well as IL12p40 associated with persistence; rs1518110 and IL20 (rs5030737), 54 (rs1800450) and 57 (rs1800451) associated with persistence</td>
</tr>
<tr>
<td></td>
<td>TNFA (1SNP)</td>
<td>Gambia, case-control (N=507)</td>
<td>−308G/A (rs1800629) associated with persistence</td>
</tr>
<tr>
<td></td>
<td>VDR (1SNP)</td>
<td>Gambia, case-control (N=530)</td>
<td>TaqI SNP (rs731236) associated with resistance to infection</td>
</tr>
<tr>
<td></td>
<td>MBL2 [also known as MBP] (3 SNPs)</td>
<td>Gambia, case-control (N=337)</td>
<td>No association with SNP (rs5030737), 54 (rs1800451) and 57 (rs1800450)</td>
</tr>
<tr>
<td></td>
<td>CTLA4 (6 SNPs)</td>
<td>USA, case-control (N=527 including 22% African-Americans)</td>
<td>−1722C/+49A (rs733618/rs3087243) correlated with persistence</td>
</tr>
<tr>
<td></td>
<td>GSTM1 (deletion), GSTT1 (deletion), EPHX1 (1 SNP), XRCC1 (1 SNP)</td>
<td>The Gambia, case-control (N=624)</td>
<td>GSTM1 deletion and XRCC1 R399Q, rs25487 associated with aflatoxin-related combined liver disease</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Gambia, case-control (N=357)</td>
<td>Increased mean adduct levels in infected with GSTM1 null, EPHX1 H113Y and XRCC1 heterozygotes at +6230A (rs3087243)</td>
</tr>
<tr>
<td>Marker</td>
<td>Description</td>
<td>Population</td>
<td>Study Details</td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
<td>------------</td>
<td>---------------</td>
</tr>
<tr>
<td>GSTM1 (deletion), GSTT1 (deletion), EPHX1 (2 SNPs)</td>
<td></td>
<td>Ghana (N=114)</td>
<td>Female gender, HBV status and EPHX1 R139H (rs2234922) associated with increased aflatoxin-albumin adducts (i.e. risk factor for HCC in individuals with GSTM1 null genotype)</td>
</tr>
<tr>
<td>GSTM1 (deletion), GSTT1 (deletion), EPHX1 (2 SNPs)</td>
<td></td>
<td>Sudan, case-control (N=355)</td>
<td>Correlations between aflatoxin exposure and HCC in individuals with the GSTM1 null genotype</td>
</tr>
<tr>
<td>SOD2 [also known as Mn-SOD] (1 SNP)</td>
<td></td>
<td>Morocco, case-control (N=318 including 19 HbsAg positives and 111 anti-HCV positives)</td>
<td>A16V (rs4880; previously associated with HCC, increased risk in HCV infecteds (Note: HBV infecteds not analysed separately, but higher frequency of HCC in HBsAg +))</td>
</tr>
<tr>
<td>IL2, IL4, IL6, IL10, IL12B, HLA-A, -B, -Cw, DRB1 DRQ1</td>
<td></td>
<td>USA, case control (including N=104 African-Americans)</td>
<td>HLA-Cw<em>03, DRB1</em>07, DQB1<em>02 (as well as related haplotypes) and IL12B allele 2 (4pb deletion) associated with non-responder phenotype; DRB1</em>15, DQB1*06 and IL12B (rs2243248, rs2243250, rs1883276) with response</td>
</tr>
<tr>
<td>HLA DRB1</td>
<td></td>
<td>Gambia, twins (N=174 pairs)</td>
<td>Heritability for anti-HBs variation due to DRB1 loci</td>
</tr>
<tr>
<td>133 mostly non-HLA genes (715 SNPs)</td>
<td></td>
<td>Gambia, cohort/case-control (N=662+393 as 2-stage screen)</td>
<td>IFNG (rs2069727), CD58 (rs1016140), MAPK8 (rs3827680, rs10857565), IL10RA (rs2508450, rs2287047), IL19 (rs12409415-rs2056225-rs2287047-rs997049-rs353644-rs7937602), ITGAL (R719V, rs4243232) and haplotypes [rs1414275-rs11588376-rs353644-rs353630-rs7937602-rs2056974-rs2473316], [rs12409415-rs2056225-rs2287047-rs997049-rs353630-rs7937602] with peak anti-HBs level; CD163 (rs6488340, rs4880) anti-HBc positivity despite vaccination</td>
</tr>
</tbody>
</table>

Note: References in this table are limited to studies relating to African or African-American populations, reports relating to other populations are summarised in the main text of this review.
Figure captions

Figure 1: Prevalence of HBsAg in Africa.

Figure 2: Hepatitis B vaccination coverage, WHO/UNICEF estimates 1980-2006
Figure 1: Prevalence of HBsAg in Africa.


Permission to reproduce granted by Cambridge University Press.
Figure 2: Hepatitis B vaccination coverage, WHO/UNICEF estimates

WHO disclaimer: "These maps do not imply the expression of any opinion whatsoever on the part of the World Health Organization concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries."