Abstract

In 2016, the global health sector strategy (GHSS) on viral hepatitis called for elimination of hepatitis B as a major public health threat by 2030 (i.e., 90% reduction in incidence and 65% in mortality). But persistence or clearance of hepatitis B virus (HBV) infection mainly depends upon host immune responses. The human leukocyte antigen (HLA) system is the center of host immune responses. HLA genes are located in chromosome 6p21.31 and cover 0.13% of the human genome and show a high degree of polymorphism and extensive patterns of linkage disequilibrium (LD), which differ among populations. The HLA genes include HLA class I, HLA class II, and other non-HLA alleles. HLA class II gene polymorphisms are strongly associated with not only persistent HBV infection but also spontaneous HBV clearance and seroconversion, disease progression, and the development of liver cirrhosis (LC) and HBV-related hepatocellular carcinoma (HCC) in chronic hepatitis B. This chapter summarizes the reported associations of HLA class II gene polymorphisms with the outcomes of HBV infection and their related mechanisms.

Keywords: HBV, HBV infection, HLA, HLA class II gene polymorphisms

1. Introduction

Since the discovery of hepatitis B virus (HBV) by Blumberg et al. in 1965, much has been elucidated regarding its virion, infection, prevention, and treatment [1, 2]. However, HBV infection continues to be a significant public health problem that has not yet been fully addressed worldwide [3, 4].

HBV infection is caused by HBV, an enveloped DNA virus that infects the liver causing hepatocellular inflammation and necrosis. HBV infection can be either acute or chronic, and...
the associated illness ranges in severity from asymptomatic to symptomatic and progressive
disease. Chronic hepatitis B (CHB) is defined as persistence of hepatitis B surface antigen
(HBsAg) for 6 months or more [2]. The World Health Organization (WHO) reports that
80–90% of infants infected during the first year of life and 30–50% of children infected before
the age of 6 years develop chronic infections; less than 5% of otherwise healthy persons who
are infected as adults will develop chronic infection; and 20–30% of adults who are chroni-
cally infected will develop cirrhosis and/or liver cancer [4]. In 2015, 257 million people lived
with HBV infection defined as hepatitis B surface antigen positive; and hepatitis B resulted in
887,000 deaths, mostly from complications including liver cirrhosis (LC) and hepatocellular
carcinoma (HCC) [4]. Once infected with HBV, one of the main risks is the development of
cirrhosis, hepatic decompensation, and ultimately HCC [2–6]. In 2016, the global health sector
strategy (GHSS) on viral hepatitis called for elimination of viral hepatitis as a major public
health threat by 2030 (i.e., 90% reduction in incidence and 65% in mortality) [5]. In 2017,
WHO’s first-ever Global hepatitis report presented the baseline values for each of the core
indicators of the strategy [6].

The prevention component of elimination is on track with respect to hepatitis B vaccination,
blood safety, and injection safety [7–10]. A promising but limited start in hepatitis testing
and treatment needs to be followed by immediate and sustained action so that we reach the
service coverage targets required to achieve elimination by 2030 [3]. Levrero et al. [11], Vyas
et al. [12], and Yoo et al. [13] reviewed that multiple emerging drug therapies are currently in
the early stages of development as part of the growing effort to find a true cure for HBV. But
persistence or clearance of HBV infection mainly depends upon host immune responses.

The major histocompatibility complex (MHC) was discovered in the mouse in 1936 [14].
The human leukocyte antigen (HLA) system, MHC in humans, is the center of host immune
responses [15–17]. HLA genes are located in chromosome 6p21.31 [18] and cover 0.13% of the
human genome [19] and show a high degree of polymorphism and extensive patterns of link-
age disequilibrium (LD), which differ among populations. The HLA genes include HLA class I,
HLA class II, and other non-HLA alleles [20]. HLA class I alleles include the three classic HLA
gene loci: HLA-A, HLA-B, and HLA-C; three non-classic HLA-E, HLA-F, and HLA-G gene
loci, which show limited polymorphism compared to the classic class I loci; and other related
non-coding genes and pseudogenes [20]. The main function of HLA class I molecules, which
are expressed in all nucleated cells, is to present non-self antigens derived from intracellular
sources, such as viruses, to CD8+ T cells (cytotoxic T cells, CTL), which then identify and kill
infected cells [21]. CD8+ T cells interact with the cognate peptide-MHC I complexes via their
T-cell receptor (TCR) and co-receptor molecule CD8. HLA class II alleles include the classic
gene loci HLA-DR, HLA-DQ, and HLA-DP and also the non-classic HLA-DO and HLA-DM
loci [20]. The classic genes are expressed on the surface of professional antigen-presenting
cells (APCs), which take up antigens derived from extracellular sources [22], such as bacteria
or food, and present them to CD4+ T helper cells. This leads to the secretion of various small
proteins, including cytokines, which regulate other immune cells such as macrophages or B
cells. In turn, macrophages can destroy ingested microbes, and activated B cells can secrete
antibodies. CD4+ T cells interact with the cognate peptide-MHC II complexes via their TCR
and the co-receptor molecule CD4. Non-classic molecules are exposed in internal membranes
in lysosomes, which help load antigenic peptides on to classic MHC class II molecules. Over the past 50 years, polymorphisms in the HLA locus have been shown to influence many critical biological traits and individuals’ susceptibility to complex, autoimmune, and infectious diseases [17, 23, 24]. Since 1979, Kew et al. [25] started the research for the association between histocompatibility antigens and HBV infection, and a plenty of researches demonstrated that the highly polymorphic HLA classes I and II genes can affect the ability of HLA molecules to trigger immune responses, which affects the outcomes of HBV infection, and discrepant conclusions were reached in different cohorts [26, 27]. HLA class II gene polymorphisms are strongly associated with not only persistent HBV infection but also spontaneous HBV clearance and seroconversion, disease progression, and the development of LC and HBV-related HCC in chronic hepatitis B [28–30]. This chapter summarizes the reported associations of HLA class II gene polymorphisms with susceptibility to HBV infection, resolution, and disease progression and their related mechanisms.

2. HLA class II gene and the clinical outcomes of HBV infection

HLA class II gene includes HLA-DRA, -DRB, -DQB, -DPA, -DPB, -DOA, -DOB, -DMA, and -DMB with 4857 alleles known with the latest report update on April 16, 2018 (http://www.ebi.ac.uk/imgt/hla/stats.html). HLA-DRB1 has the most allelic variability with 2165 alleles, and in turn HLA-DQB1 with 1196 alleles, HLA-DPB1 with 975 alleles, and HLA-DRB3 with 157 alleles [20].

2.1. HLA class II allele polymorphism and HBV infection outcomes

HLA-DR is widely used in transplant gene matching [31] and as an activated T cell surface marker [32], and early related to the clinical outcomes of HBV infection [33]. In 2013, we reviewed the relationship between HLA class II alleles and HBV infection, there are the most number of HLA-DR alleles related to HBV infection, such as HLA-DRB*03, 07, 09 and 12, may be the risk factors of HBV infection; and HLA-DRB*04, 11, 13 and 14 may be the protect factors of HBV infection [26]. In 2016, Wang et al. also reviewed the relationship of HLA-DR alleles with HBV infection [27], including HLA-DRB*1301-02 which is consistently associated with HBV clearance globally, such as in Gambia, Germany, Korea, and Spain; HLA-DRB*11 and -DRB*14 are associated with spontaneous recovery in patients with HBV subgenotype C2 infection in Northeast China; HLA-DR*0406, and HLA-DR, antigens are associated with protective effect on acute HBV infection; and HLA-DRB*08 and -DRB*09 alleles, which are susceptible to HBV infection, were found in Brazilian populations determined in young and male blood donors. Meanwhile, HLA-DRB*11/12 alleles are associated with HBV persistence globally, and HLA-DRB*07 is the only one associated with infant susceptibility to intrauterine HBV infection and a significant negative predictor of cirrhosis. Our researches also showed that -DRB*13 may protect subjects from HBV infection [26]; -DRB*12 may have a high risk for HBV infection [26]; and -DRB*07 and 12 may be implied in viral persistence [26, 34]. Analysis still identified HLA-DRB*12:02 as the top susceptible HLA allele associated with acute-on-chronic liver failure (ACLF) [35]. A large number of studies have been
conducted to identify HLA-DRB1 genetic variants associated with HCC risk and clinical outcomes, but many of the findings in these studies are inconsistent and inconclusive [36–38]. A meta-analysis by Lin et al. reported an ethnicity-dependent association between specific HLA-DRB1 alleles and HCC risk, DRB1*07 and DRB1*12 were significantly associated with the risk of HCC in the whole populations, and DRB1*15 allele significantly increased the risk of hepatocellular carcinoma only in Asians [36]. One research by Ma et al. suggests that if genetic factors play a role in familial aggregation of hepatocellular carcinoma, the deficiency in the DRB1*11 and DRB1*12 alleles might be the risk factor at work in the Guangxi Zhuang Autonomous Region, P.R.C. [37]. Another meta-analysis by Liu et al. reported that HLA-DRB1*01 and 11 alleles were protective factors, while HLA-DRB1*12 and 14 alleles were risk factors for HCC development [38]. These findings are somewhat reasonable considering the incidence and distribution of HCC which are closely linked to environmental, dietary, and lifestyle factors, as well as genetic profiles, but the risk for HCC was not controlled for possible confounders such as HBV or HCV, which are more prone to bias than that of the randomized clinical trial studies.

HLA-DQ belongs to one of HLA class II molecules; it is also expressed as cell-surface glycoproteins that bind to exogenous antigens and present them to CD4+ T cells. HLA-DQ molecules function as a heterodimer of alpha and beta subunits which are encoded by the HLA-DQA1 and HLA-DQB1 genes, respectively. HLA-DQs are highly polymorphic especially in exon 2 which encode antigen-binding sites. Therefore, a number of alleles have been declared to be associated with persistent HBV infection [26, 27]—HLA-DQA1*0102, 0201, 0301, and 0402 and HLA-DQB1*0604, and so on associated with HBV clearance; and HLA-DQA1*0103, 0201, 0302, and 0501 and HLA DQB1*0301, and so on associated with HBV persistence [26]; the two-locus haplotype consisting of -DQA1*0501 and -DQB1*0301, and the three-locus haplotype consisting of -DQA1*0501, -DQB1*0301, and -DRB1*1102 were significantly associated with persistent HBV infection in an African-American cohort; -DQB1*0301 was associated with HBV persistence globally; in addition, -DQB1*0201 is a HBV-resistant gene, and -DQB1*0303 is a susceptibility gene of carrying HBV in Xinjiang Uygur ethnic groups of China; and -DQB1*0503 are associated with early HBeAg seroconversion in CHB children in Taiwan [27]. HLA genotyping-based analysis identified -DQB1*0601 as having the strongest association, showing a greater association with CHB susceptibility [28].

The HLA DPA1 and HLA DPB1 belong to the HLA class II alpha and beta chain paralogs, which also make a heterodimer consisting of an alpha and a beta chain on the surface of antigen-presenting cells. This HLA class II molecule also plays a central role in the immune system by presenting peptides derived from extracellular proteins. Identification of a total of five alleles, including two risk alleles (DPB1*09:01 and DPB1*05:01) and three protective alleles (DPB1*04:01, DPB1*04:02, and DPB1*02:01), would enable HBV-infected individuals to be classified into groups according to the treatment requirements. Moreover, among the five reported HLA-DPB1 susceptibility alleles, three DPB1 alleles (DPB1*05:01, *02:01, and *04:02) had primary effects on CHB susceptibility. However, the association of the remaining two alleles (DPB1*09:01 and *04:01) had come from LD with HLA-DR-DQ haplotypes (i.e., DRB1*15:02-DQB1*06:01 and DRB1*13:02-DQB1*06:04, respectively) [28, 39].
2.2. Single nucleotide polymorphisms at HLA class II gene and HBV infection outcomes

HLA gene variations are strongly associated with HBV infection outcomes in not only HLA alleles but also single nucleotide polymorphisms (SNPs) identified through genome-wide associated studies (GWASs). Recent GWASs have revealed several SNPs at HLA class II region associated with the risk of HBV infection [23, 30].

A Chinese study by Zhu et al. [40] identified two HLA-DR loci that independently drive chronic HBV infection, including HLA-DRβ113 sites 71 and rs400488. Acute-on-chronic liver failure (ACLF) is an extreme condition after severe acute exacerbation of chronic hepatitis B. Tan et al. carried out a genome-wide association study, among 1300 ACLFs and 2087 AsCs, and identified rs3129859 at HLA class II region (chromosome 6p21.32) which is associated with HBV-related ACLF. Analysis identified HLA-DRB1*12:02 as the top susceptible HLA allele associated with ACLF. The association of rs3129859 was robust in ACLF subgroups or HBV e antigen-negative chronic hepatitis B phase. Clinical traits analysis in patients with ACLF showed that the risky rs3129859 C allele was also associated with prolonged prothrombin time, faster progression to ascites development, and higher 28-day mortality [35]. SNP rs9272105 locates between HLA-DQA1 and HLA-DRB1 on 6p21.32. SNP imputation in the GWAS discovery samples revealed additional SNPs showing association, but rs9272105 remained the top SNP within the region [41], which successfully validated the associations between rs9272105 and HCC risk [42]. Of the 12 SNPs reported in HBV-related HCC GWASs, rs2647073 and rs9997872 near HLA-DRB1 were found to be significantly associated with the risk of HBV-related LC, which suggested that genetic variants associated with HBV-related hepatocarcinogenesis may already play an important role in the progression from CHB to LC [43]. A recent study reported new SNPs at HLA-DRB1 (rs35445101) associated with TP53 expression status in HBV-related hepatocellular carcinoma [30].

In 2013, Jiang et al. first found the association of HCC risk with rs9275319 at 6p21.3 located between HLA-DQβ and HLA-DQA1, which was not reported in earlier GWASs of HCC [44]. Their following research and that of Wen et al. successfully validated the associations between rs9275319 and HCC risk [42, 43]. Three SNPs belonging to the HLA-DQ region (rs2856718, rs7453920, and rs9275572) were studied. HLA-DQ rs2856718G, rs7453920A, and rs9275572A were strongly associated with decreased risk of chronic HBV infection and natural clearance; while rs2856718A, rs7453920G, and rs9275572G served as a risk factor in HBV infection in Japanese populations and in Southeast China [45–48]. Chang et al. found that rs9276370 (HLA-DQA1), rs7756316, and rs7453920 (HLA-DQB1) are significantly associated with persistent HBV infection, especially the “T-T” haplotype composed of rs7756316 and rs9276370 that is more prevalent in severe disease subgroups and associated with nonsustained therapeutic response in male Taiwan Han Chinese individuals [49]. A nearest study reported four new SNPs at HLA-DQB1 (rs1130399, rs1049056, rs1049059, and rs1049060) associated with TP53 expression status in HBV-related HCC [30].

A Chinese study by Zhu et al. [40] identified HLA-DPβ1 positions 84–87 that independently drive chronic HBV infection. In 2009, Katamani et al. first reported two SNPs with
the strongest relation to HBV infection from the HLA-DP locus: rs3077 on HLA DPA1 and rs9277535 on HLA DPB1 in Japanese and Thai populations [50, 51]. Subsequently, plenty of studies further demonstrated their roles [27, 52–54]. rs3077 and rs9277535 are significantly related to HBV persistent infection and both A alleles of these two SNPs are protection alleles in Chinese Han [55], Japanese and Korean [56], and European [57] populations, while in Chinese Zhuang subjects, only HLA-DP rs9277535A is associated with decreased risk [55]. Also, only a highly significant association of HLA-DFA, rs3077C with HBV infection was observed in Caucasians [58]. HBeAg-negative HBV carriers with rs9277535 non-GG genotype had a higher chance to clear HBsAg. Compared to GG haplotype of rs3077 and rs9277535, GA haplotype had a higher chance of achieving spontaneous HBsAg loss in Chinese subjects of Taiwan [59]. On the whole, the present findings show that SNPs rs3077 and rs9277535 at HLA-DP locus protect against HBV infection and increase the chance of HBV clearance, while the importance of these polymorphisms as a predictor of HCC may be limited [45, 60, 61]. In a report by Hu et al., HLA-DP rs3077 showed an approaching significant effect on susceptibility to HBV persistent infection and HCC development when considering multiple testing adjustments [62]. Li et al. found evidence for the association at rs9277535 with HCC independently by imputation [41]. Thomas et al. reported that SNPs rs3077 and rs9277535 that associated most significantly with chronic hepatitis B and outcomes of HBV infection in Asians had a marginal effect on HBV recovery in European and African-American samples. However, they identified a novel variant in the HLA-DPB1 3'UTR region, 496A/G (rs9277534), which associated very significantly with HBV recovery in both European and African-American populations [63]. Hu et al. also found that the variant at rs9277534 could affect both the spontaneous clearance of HBV infection and progression from asymptomatic HBV carriers to HBV-related liver cirrhosis in southwest Han Chinese population [64]. Chang et al. found that rs9366816 near HLA-DPA3 are significantly associated with persistent HBV infection in male Taiwan Han Chinese individuals [49]. A nearest study reported a new SNP at HLA-DPB1 (rs1042153) associated with TP53 expression status in HBV-related hepatocellular carcinoma [30].

2.3. The role and mechanism of HLA class II gene variations associated with HBV infection outcomes

HLA class II genes encode proteins expressed on the surface of antigen-presenting cells such as macrophages, dendritic cells, and B cells, and thereby have a critical role in the presentation of antigens to CD4+ T-helper lymphocytes. In our previous review, there were three mechanisms related to HBV infection outcomes, including HLA molecular structure, HLA gene expression, and its regulatory [26].

HLA class II genes have many structural variants that have been linked to immune response [65, 66] to autoimmune diseases [67, 68], idiosyncratic drug toxicity (IDT) [69, 70], and infectious agents [71, 72]. But the structural variants related to HBV infection outcomes in HLA class II genes only have our previous research report [73]. We know that the HLA class II molecules are heterodimeric glycoproteins, which are able to present peptides to CD4+ T cells; the primary protein sequences (α1 and β1 chains) encoded by the second exon of the HLA class II genes comprise the peptide-binding grooves that accommodate amino acid chains of the bound peptides; the specificity of the peptide-binding grooves is governed by the properties of pockets in the
grooves which include nine different structural pockets (Ps) from P1 to P9, typical pockets being P1, P4, P6, and P9, which accommodate the antigen peptide side chains; and the polymorphic residues encoded by polymorphic HLA class II genes can influence the structural (size and shape) and electrostatic properties and further function of the pockets. So, the determination of the structural and electrostatic properties of the HLA class II peptide-binding grooves associated with diseases may help identify the disease mechanism. We found that DR07 and DR12 carry amino acids Leu and His at residue 30 of HLA-DRβ1 chain, respectively, and Val at residue 57 of HLA-DRβ1 chain, difference from Tyr30 and Asp57 carried by DR04 and DR11, leading to present positive charge at P9 [73], as well as increasing size due to the absence of an intact salt bridge at P6 and P9 [73, 74]. Hence, which HLA DR07 and DR12 is sufficient to chronic HBV infection is supported by the structure characters resulting from HLA gene polymorphism.

Gene expression, that is, the qualitative and quantitative expression of mRNA transcripts from DNA templates, forms a first link in the functional path between nucleotide sequence and higher-order organismal phenotypes. Despite the undeniable importance of a controlled gene regulatory process in which proper transcripts are expressed at the correct time and location, studies have shown that there is widespread inter-individual variability with respect to gene expression, or mRNA levels, within a given cell type or tissue [75]. Early in 1986, Edwards et al. [76] examined frozen sections of human fetal spleen from 12 to 20 weeks of gestation by using polyclonal antibodies to Ig isotypes, monoclonal antibodies to HLA class II subregion locus products, B and T cells, and follicular dendritic cells. Their data suggest that class II antigens are differentially expressed on developing lymphoid cells; DR and DP expression occurring in the earliest spleens examined, with expression of DP on a subpopulation of DR-positive cells; IgD and DQ expression appears to be coincident on maturing B cells as they begin to form follicles, and an immunoregulatory role for HLA-DQ in B cell development is implicated. In fact, HLA class II gene (HLA-DR) expression level is still one of the markers for immune reaction related to disease [33, 77]. In 2017, we reviewed the association of HLA-DR expression level with diseases and reported our studies about the characteristics of HLA-DR expression in patients with different outcomes of HBV infection. Compared to persons with no HBV marker, HBV infection and vaccination induce increased expression of HLA-DR, especially in the clearance of HBV infection [78]. Genetic variants that influence HLA mRNA expression might also affect antigen presentation and many “gene expression-associated SNPs” (eSNPs) have been found for HLA genes [79, 80]. An integrated approach combining genotype information with genome-wide gene expression data in relevant tissues can identify genetic variations that are both regulatory and disease causing [81]. Cavalli et al. believed that a majority of causal genetic variants underlying complex diseases appear to involve regulatory elements, rather than coding variations [82]. Kaur et al. reported that structural and regulatory diversity shape HLA-C protein expression levels and that quantitative variation in the expression of HLA-C can influence the clinical course of HIV infection and the risk of graft-versus-host disease [83]. By influencing HLA mRNA expression, rs3077 and rs9277535 variants, both are noncoding variation (3′UTR) in the HLA-DPA1 and HLA-DPB1 region, and are related to enhanced clearance of hepatitis B virus infection [57, 63] and increased risk of graft-versus-host disease in mismatched hematopoietic cell transplant recipients [84, 85]. These findings were identified by a study about mapping of hepatic expression quantitative trait loci (eQTLs) in a Han Chinese population by Wang et al. [86].
eQTLs, namely, the discovery of genetic variants, explain variation in gene expression levels which has significant differences in mean expression levels between population, tissue or cell type [75, 87]. Typically, in the eQTL mapping literature, regulatory variants have been characterized as either cis or trans acting, reflecting the predicted nature of interactions and of course depending on the physical distance from the gene they regulate. Conventionally, variants within 1–2 Mb (megabase) on either side of a gene’s TSS were called cis, while those at least 5 Mb downstream or upstream of the TSS or on a different chromosome were considered trans acting [87]. The most significant GWAS SNPs are strong eQTLs and proposing candidate disease genes. Multiple variants of low-effect sizes affect multiple genes by gene regulatory network. A network component can be viewed as a cis effect that transmits its signal in trans essentially making the cis SNP a trans SNP as well [87]. Such studies have offered promise not just for the characterization of functional sequence variation but also for the understanding of basic processes of gene regulation and interpretation of genome-wide association studies [75, 87].

Up to now, many research findings have demonstrated that HLA class II gene variations, including allele polymorphisms and SNPs, are associated with HBV infection outcomes by influencing the molecular structure and the expression level of HLA class II molecules expressed on the cell surface. As the researches progress with methodological improvements, such as the prediction of gene expression [88] and PrediXcan analysis [89], more structure variations, cis SNPs and trans SNPs, were found underlying this disease. This can help understand the mechanisms linking genome-wide association loci to the disease, and implement precise individualized prevention, diagnosis, and treatment of the disease.

Acknowledgements

This work was supported by Wu Jieping Medical Foundation [Grant number 320.6750.18230].

Conflict of interest

None.

Author details

Shuyun Zhang

Address all correspondence to: zhang13214501198@163.com

Research Center, The Second Affiliated Hospital of Harbin Medical University, Harbin, Heilongjiang Province, P.R. China

References


[33] van den Oord JJ, de Vos R, Desmet VJ. In situ distribution of major histocompatibility complex products and viral antigens in chronic hepatitis B virus infection: Evidence that HBC-containing hepatocytes may express HLA-DR antigens. Hepatology. 1986;6(5):981-989. PMID: 3530950


[71] Delgado JC, Baena A, Thim S, Goldfeld AE. Aspartic acid homozygosity at codon 57 of HLA-DQ beta is associated with susceptibility to pulmonary tuberculosis in Cambodia. Journal of Immunology. 2006;176(2):1090-1097. DOI: https://doi.org/10.4049/jimmunol.176.2.1090


[89] Zeng P, Wang T, Huang S. Cis-SNPs set testing and predicxcan analysis for gene expression data using linear mixed models. Scientific Reports. 2017;7(1):15237. DOI: 10.1038/s41598-017-15055-8