Hepatitis B Virus Genetic Diversity: Disease Pathogenesis

Maria Kuttikan Jayalakshmi, Narayanan Kalyanaraman and Ramasamy Pitchappan

1. Introduction

Hepatitis B Virus (HBV) infection is a global health problem: an estimated two billion people (one-third of the global population) have been infected with HBV at some point in their life; of these, more than 350 million suffer from chronic HBV infection, resulting in over 600,000 deaths each year, mainly from cirrhosis or liver cancer [1]. More than 10% of the global chronic HBV population resides in India [2]; infection may lead to liver damage that results in acute or chronic hepatitis, liver cirrhosis, and hepatocellular carcinoma (HCC) (Figure 1). HBV infection was first identified in 1965 when Blumberg and co-workers [3] found the hepatitis B surface antigen (HBsAg), originally termed as Australia antigen. Enhanced viral replication leading to a vigorous and extensive immune response may lead to massive liver injury resulting spontaneously into fulminant hepatic failure. The seriousness of disease incidence is mainly related to various host factors (age, gender, duration of infection, immune response) and viral factors (viral load, genotype, quasispecies) (Figure 2). Recent evidence shows that considerable molecular variation occurs throughout the HBV genome, which is correlated with geographical distribution of genotypes and severity of disease.

Figure 1. HBV mediated liver damage among various stages
1.1. HBV evolution

The history and origin of HBV is partially understood. The evolution of HBV can be traced from lowest vertebrates such as the birds infecting virus i.e. avian hepadnavirus which shares a sequence homology of 40%, non-primate such as rodents infecting wood chuck HBV viruses shares 80% homology [4, 5]. However, highest incidence of greater than 94% of homology is found to occur in primates infecting viruses suggesting an evolutionary sequence for the human HBV [6].

The long time evolution of HBV therefore leads to the occurrence of various genotypes, subgenotypes, mutants, recombinants and even quasispecies [7]. Major forces such as genetic drift, bottle neck effects, founder effects and recombination played a vital role in evolution, adaptation of HBV and become a successful pathogen in the host. The genetic variability, in conjunction with the migration of human race, has led to the divergence of HBV into genetically different groups, called genotypes, with a distinct geographic distribution. Some researchers suggest that HBV co-evolved with modern humans, as they migrated from Africa, around 100,000 years ago [8,9]. Based on the prevalence, geographic distribution and characteristic of recombinant genotypes, it can be suggested that genotypes A and D co-exist over a relatively long period, while genotypes B and C are believed to have recent epidemiological contact in Asia [10]. The genetic diversity of HBV and its geographical distribution may help us to reconstruct the evolutionary trend of HBV. This may help to generate additional genetic data on the evolution and migration pattern of man [8].

1.2. Hepatitis B Virus: Molecular virology

HBV is a hepatotropic, non-cytopathic virus and a prototype member of the family Hepadnaviridae with a genome size of ~3,200 base pairs. The viral genome consists of a partially double-stranded, relaxed-circular DNA (RC-DNA), comprising a complete coding strand (negative strand) and an incomplete non-coding strand (positive strand), which replicates by reverse transcription via an RNA intermediate. Due to reduced fidelity of the reverse transcription process, this pregenomic RNA (pgRNA) is prone to mutation. The genome encodes four overlapping reading frames that are translated to make the viral core protein (HBcAg), the surface proteins (HBsAg), a reverse transcriptase (RT), and the hepatitis B “x” antigen (HBxAg).
1.3. Viral entry and replication

HBV has a high degree of species and tissue specificity that results in very high levels of viral replication without actually killing the infected cell directly. The mechanism through which HBV enters hepatocyte or other susceptible cells remains elusive mostly owing to lack of a proper cell culture system. As a pararetrovirus, HBV uses reverse transcription to copy its DNA genome and lack of proof-reading capability permits the emergence of mutant viral genomes and quasispecies.

Upon infecting a hepatocyte, the HBV genome is delivered into the nuclear compartment where cellular repair enzymes are involved in repairing the viral genome into covalently closed circular DNA (cccDNA). This viral DNA acts as a transcriptional template [11, 12] for the generation of the pregenomic mRNA (pg mRNA), pre-core mRNA and all other subviral mRNAs [13, 14]. Subsequently, cccDNA is chromatinized into viral minichromosome that ultimately serves as an intrahepatic reservoir of HBV and stays throughout the life of the chronically infected host [15, 16]. The pregenomic RNA is encapsulated by the virion core particle and reversely transcribed by the viral polymerase, forming a single-strand DNA (negative strand). Subsequently, the pregenome is degraded and the negative strand DNA then acts as a template for synthesis of a positive strand DNA with variable length. Finally, the HBV genome is either encapsulated to produce virions to be secreted out, or recycled back to the nucleus to maintain a pool of cccDNA, resulting in the formation of a steady-state population of 5–50 copies of cccDNA molecules per infected hepatocyte [14, 17] (Figure 3).

Figure 3. Viral entry and replication cycle of HBV

2. HBV genomic organization and proteins

The HBV genomes comprise a partially double-stranded 3.2kb DNA molecule, organized into 4 overlapping open-reading frames (ORFs). Four sets of mRNAs are then transcribed
from the viral minichromosomes using host cell machinery, RNA polymerase II. These molecules are then transported by cellular proteins to the cytoplasm where they are translated to produce the viral proteins: hepatitis B core antigen (HBcAg or nucleocapsid protein from the 3.5 kb RNA); the soluble and secreted hepatitis B e antigen (HBeAg, from the 3.5 kb RNA); the Pol protein (from the 3.5 kb RNA, which is longer than the viral entire genome); the viral envelope proteins HBsAg (from the 2.4 and 2.0 kb RNAs) and hepatitis B X protein (HBx from the 0.7 kb RNA). The dynamic nature of individual viral proteins and insufficient immune response elicited by the infected host immune cells lead to the persistence of HBV infection.

![Hepatitis B Virus genome organization](image)

**Figure 4.**

### 2.1. Hepatitis B virus surface antigen (HBsAg)

Since Blumberg’s discovery of HBsAg in 1965, it has been used as the hallmark for the diagnosis of HBV infection [18]. HBsAg is the prototype serological marker of HBV infection that characteristically appears after 1 to 10 weeks of an acute exposure to HBV but before the onset of visible symptoms or elevation of serum alanine aminotransferase (ALT) [19]. HBsAg circulates in a wide array of particulate forms such as competent virions (42 nm, Dane particles), 20 nm diameter filaments of variable length, and 20–22 nm spherical defective particles, corresponding to empty viral envelopes. It exceeds virions by a variable factor of $10^2$ – $10^5$ and accumulates several hundred micrograms per ml of serum [13]. The principal function of the HBs protein as a virological structure is to enclose the viral components, in addition to playing a major role in cell membrane attachment to initiate the infection process by binding to the hepatocyte plasma membrane [20]. Persistence of HBsAg for more than 6 months indicates chronic infection and it is estimated that fewer than 5% of immunocompetent adult patients with acute hepatitis B progress to chronic infection [21]. The immune response enhancing ability of HBsAg is not clear yet it is understood that large amounts of HBsAg may induce T cell anergy, leading to decreased antibody-mediated neutralization of HBV and generalized hyporesponsiveness towards pathogens.
2.2. Hepatitis B core antigen (HBcAg)

HBcAg is the major constituent of the nucleocapsid, which is essential for viral replication. It also forms a part of icosahedral subviral particles that packs the viral reverse polymerase and the pregenome [22] derived from the ORF-C. It has either 183 or 185 amino acids depending on the genotype of the virus. HBcAg is a particulate and multivalent protein antigen which can function as both a T cell-independent and a T cell-dependent Ag [23] and is ~1000 fold more immunogenic than the HBeAg [24]. The response of T cells to HBcAg has been reported to contribute to the resolution and seroconversion in chronic hepatitis B [25].

2.3. Hepatitis B envelope antigen (HBeAg)

HBeAg is an accessory protein of HBV, not essential for replication in vivo [26, 27] but important for natural infection. This antigen has been used clinically as an index of viral replication, infectivity, severity of disease, and response to treatment [28]. HBeAg may play a role in perpetuating viral infection during perinatal transmission, often resulting in chronic infection and eliciting HBe/HBcAg-specific T helper cell tolerance in utero [29, 30].

HBeAg is a non-particulate secretory protein discovered by Magnus and Espmark in 1972 [31]. It is derived after cleavage from a 212 amino acid precursor, precore protein that is encoded by the HBV precore gene (the pre-C sequence and C-gene). It is highly conserved evolutionarily between all Hepadnaviridae. As part of the core protein, it has a nucleic acid binding activity [32] required for the pregenomic RNA encapsidation, and modulation of polymerase activity for reverse transcription of pregenome [33,34].

HBeAg is found at concentrations of greater than 10μg mL⁻¹ in the plasma, which can be detected even by agar-gel immunodiffusion. Secreted HBeAg has an immunoregulatory function in utero by establishing T cell tolerance to HBeAg and HBcAg, which may predispose neonates born to HBV-infected mothers to develop persistent HBV infection [29]. Milich et al., [35-37] further demonstrated an immunomodulatory role of HBeAg in antigen presentation and recognition by CD4⁺ cells.

2.4. Hepatitis B X antigen (HBxAg)

HBx antigen, a 17 kDa multifunctional, non-structural protein, comprised of 154 amino acids, which is conserved across all the mammalian infecting Hepadnaviridae. The X gene is the smallest of the four partially overlapping ORF of the HBV genome. The biological function of HBx protein is not yet clear however this has been implicated in causing HBV associated liver cancer.

Accumulating evidences indicate that the HBV X gene is indispensable to HBV replication, propagation and integration of viral DNA into the host’s genome. The expression of full-length HBx protein is dispensable for virus production in vitro and a critical component of the infectivity process in vivo [38]. HBx protein promotes virus gene expression and replication by trans-activating the virus promoters and enhancer/promoter complexes [39,
In addition, HBx accumulation enhances viral replication by altering various cellular activities including aberrant expression of molecules, involved in host cell signal transduction, transcription and proliferation, leading to viral persistence and hepatocarcinogenesis [41, 42, 43].

3. HBV viral load

HBV is reported to be present in the blood of HbeAg seropositive individuals at a concentration of approximately $10^8$–$10^9$ viral particles mL$^{-1}$ of blood [44]. HBV DNA is present in high titers in blood and exudates of acute as well as chronic cases. Generally, moderate viral titers are found in saliva, semen and vaginal secretions [45, 46]. The serum levels of HBV DNA largely depend on the viral genotype, and the quantity of HBeAg in serum, which determines the progression of liver cirrhosis to carcinoma.

3.1. HBV genotypes and subgenotypes

The global HBV genome diversity is influenced by both genotypic and phenotypic variability; genotypes often evolve in the absence of selective pressure but phenotypic variability often develops in the presence of selective pressure exerted by host immune system or even during certain therapeutic measures [47].

This genetic diversity of HBV has been associated with differences in clinical and virological characteristics, indicating that they may play a role in the virus–host relationship [48]. Genotypes may result from neutral evolutionary drift of the virus genome, from recombination, or as a consequence of a long-term adaptation of HBV to genetic determinants of specific host populations. Structural and functional differences between genotypes can influence the severity, course and likelihood of complications, HBeAg seroconversion and response to treatment of HBV infection and possibly the vaccination against the virus [49].

Traditionally HBV was classified into 4 subtypes or serotypes (adr, adw, ayr, and ayw) based on antigenic determinants of HBsAg [50]. In the advent of more molecular approaches, serotyping of viral strains was replaced by various genotyping methods. Galibert et al. [51], published the first sequence of a complete HBV genome. Later, Okamoto et al. [52], analyzed 18 full length genomes and divided them into four groups or genotypes, named as A to D. The ability of HBV to adapt to the host genetics as well as immunogenic environment by genetic variation, led to the evolution of eight established genotypes (A-H): [8, 53] and two putative genotypes (I and J), each corresponding to a rather well-defined geographical distribution (Table 1). HBV genotypes A and D have worldwide distribution, whereas genotypes B and C are mostly found in Asia. In India HBV genotypes A and D are common in various parts, followed by genotype C specifically in eastern part of India [54-56]. The new genotype I is a complex recombinant form of genotypes A, C, and G [57,58]. The genotype J which was positioned phylogenetically in between the human and ape genotypes and was isolated from a 88-year-old hepatitis patient living in Okinawa, Japan.
who had a history of residing in Borneo during the World War II [59]. In addition, countries in which multiple genotypes circulate, co-infections and recombination events may occur leading to the emergence of hybrid strains that can become the dominant subgenotype prevailing in certain geographical regions.

Table 1. Overview of HBV major genotypes and its distribution pattern

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Serotype</th>
<th>Geography distribution</th>
<th>Genome (base pair)</th>
<th>ORF differences</th>
<th>HBV proteins (aa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A (A1 - A6)</td>
<td>adw2</td>
<td>Africa, Europe, USA, Australia</td>
<td>3221</td>
<td>Insertion of aa 153 and 154 in HBc</td>
<td>119 845 185</td>
</tr>
<tr>
<td>B (B1 - B9)</td>
<td>adw2, ayw1</td>
<td>South East Asia, China, Japan</td>
<td>3215</td>
<td>119 843 183</td>
<td></td>
</tr>
<tr>
<td>C (C1 - C16)</td>
<td>adw2, ayr, adrq-, adrq+</td>
<td>South East Asia, China, Korea, Japan, Polynesia, Australia</td>
<td>3215</td>
<td>119 843 183</td>
<td></td>
</tr>
<tr>
<td>D (D1 - D7)</td>
<td>ayw2,3 and 4</td>
<td>Mediterranean area, Middle East, East Europe, India</td>
<td>3182</td>
<td>Deletion of aa 1-11 in preS1</td>
<td>108 832 183</td>
</tr>
<tr>
<td>E (ND)</td>
<td>ayw4</td>
<td>West and Central Africa</td>
<td>3121</td>
<td>Deletion of aa 11 in preS1</td>
<td>118 842 183</td>
</tr>
<tr>
<td>F (F1 - F4)</td>
<td>adw4q-</td>
<td>South America, Central America, Alaska, Polynesia</td>
<td>3215</td>
<td>119 843 183</td>
<td></td>
</tr>
<tr>
<td>G (ND)</td>
<td>adw2</td>
<td>Europe, North America (Coinfection almost universal, with genotype A) Central America, Mexico, South</td>
<td>3248</td>
<td>Insertion of 12 aa in HBc</td>
<td>118 842 195</td>
</tr>
<tr>
<td>H (ND)</td>
<td>adw4</td>
<td>United States</td>
<td>3215</td>
<td>Deletion of aa 11 in preS1</td>
<td>119 843 183</td>
</tr>
<tr>
<td>I (ND)</td>
<td>adw</td>
<td>Vietnam, Laos, India</td>
<td>3215</td>
<td>119 843 183</td>
<td></td>
</tr>
<tr>
<td>J (ND)</td>
<td>adw</td>
<td>Japan</td>
<td>3182</td>
<td>108 832 183</td>
<td></td>
</tr>
</tbody>
</table>

ND = No subgenotypes so far identified

3.2. HBV Genotyping

Currently, more than 10 different methods have been developed for HBV genotyping with variable sensitivity, specificity, turnaround time and cost. These methods include restriction fragment length polymorphisms (RFLP) [60], PCR with specific primers for single genotypes [61], multiplex-PCR for many HBV genotypes [62, 63] and on hybridization technologies [64] or real time quantification and genotyping [65] or Mass spectrometry [66]. The gold standard method for HBV genotyping is a complete genome sequencing followed by phylogenetic analysis of the sequence divergence [9, 52]. Sequence and phylogenetic analysis can also be performed on individual genes, more often in envelope (S) gene. The reliability of using individual genes or limited gene sequence will depend both on the size of the sequence analyzed and the degree of sequence homology. The HBV genotype can be determined by other methods that are based on a limited number of conserved nucleotide or amino acid differences between the genotypes. Line probe assay (LiPA) may be a suitable alternative to sequencing but it is expensive when compared to other methods such as multiplex PCR, RFLP and serotyping. Based on the sensitivity, other methods such as
microarrays, real time PCR, reverse dot blot, restriction fragment mass polymorphism (RFMP), invader assay are being used [67,68]. Identification of HBV genotypes will be useful to understand the source of infection, predict clinical outcome at individual level and monitoring the development of newer viral strains at population level.

3.3. Clinical implications of HBV genotyping

Mounting evidence shows that HBeAg seroconversion rates, HBcAg seroconversion, viremia levels, viral latency, immune escape, emergence of mutants, pathogenesis of liver disease, response and resistance to antiviral therapy are all depend on the HBV genotypes and subgenotypes. Individual or combinations of the above factors are responsible for the degree of clinical heterogeneity displayed by the infected persons [69, 70].

Earlier reports from India, where genotypes A, D are prevalent and patients infected with genotype D had relatively high degree of disease severity and develop HCC [56]. Patients with genotypes C and D have a lower response rate to interferon therapy than patients infected with genotypes A or B [71,72]. An Alaskan population study which compares the clinical virological properties of five genotypes has shown that the mean age seroconversion from HBeAg to anti-HBe is significantly lower among genotypes A, B, D and F, than compared to genotype C [73]. This may be due to the development of more mutations in the basal core promoter region of genotypes C and D compared to genotypes A and B. Genotype G appears defective, and usually occurs together with another genotype, which provides transcription factors necessary for replication. Viral load of the patients with genotype mixture is usually higher than that of those infected with unique genotype.

3.4. HBV mutant (Phenotypic variants)

Phenotypic variants emerge in response to selective pressure [47]. Development of HBV mutants are often related to the persistence of cccDNA and viral factors such as the existence of quasispecies, high rate of HBV replication, error prone RT-based life cycle and adaptive mutants with compensatory mutations. Host factors include compliance with antiviral therapy, immune response, enhanced hepatic inflammatory response and host genetic background. HBV quasispecies arise due to the average daily production of >10^{11} virions with a error rate of 1.4-3.2 x 10^{-5} nucleotide substitutions per site per year, which results in production of all possible number of different single base changes in the HBV genome [74-76]. Active replication leads to an estimated misincorporation rate of around 10^4 owing to the lack of proof-reading (3’ – 5’ exonuclease activity) [77]. The combination of a high error rate together with an increased replication rate produces as high as 10^9 mutation day^{-1} over the entire 3.2kb genome [74, 78] but, the extreme overlapping of the ORF of the HBV genome limits the possibility of all these mutations [79].

Moreover, mutations in the HBV genome seem to play a vital role in differential outcome of this infection (Table 2). Two important mutations in the HBV virus have been associated with differential outcome such as the basal core promoter (BCP) mutation and the pre-core
Hepatitis B Virus Genetic Diversity: Disease Pathogenesis

(PCR) mutation. BCP mutation is a double substitution, A1762T, G1764A, in the basal core region of HBV. It has clearly been associated with an increased risk of HCC and cirrhosis in multiple studies, both cross-sectional and prospective [80].

<table>
<thead>
<tr>
<th>HBV Genomic region</th>
<th>Molecular Effect</th>
<th>Clinical relevance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-S/S</td>
<td>pre-S1/pre-S2 mutations</td>
<td>Vaccine escape</td>
</tr>
<tr>
<td></td>
<td>S gene mutations</td>
<td>Immune escape</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diagnostic escape</td>
</tr>
<tr>
<td></td>
<td></td>
<td>High risk for HCC</td>
</tr>
<tr>
<td>Pre-C/C</td>
<td>pre-C stop codon mutations</td>
<td>HBeAg negativity</td>
</tr>
<tr>
<td></td>
<td>Core promoter mutations</td>
<td>Viral persistence</td>
</tr>
<tr>
<td></td>
<td></td>
<td>High risk for HCC</td>
</tr>
<tr>
<td>Pol</td>
<td>Pol gene mutations</td>
<td>Viral latency</td>
</tr>
<tr>
<td></td>
<td>Mutations in YMDD motif</td>
<td>Viral persistence</td>
</tr>
<tr>
<td>Regulatory sequences</td>
<td>Core promoter mutations or Hepatocyte Nuclear Factor 1(HNF1)</td>
<td>HBeAg negativity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>High viremia</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Severe hepatitis</td>
</tr>
<tr>
<td>X gene</td>
<td>Truncated X gene (8 base pair deletion at the 3’ end)</td>
<td>Immune escape</td>
</tr>
<tr>
<td></td>
<td></td>
<td>High risk for HCC</td>
</tr>
</tbody>
</table>

Table 2. HBV mutants and its clinical impact

4. Treatment

Antiviral medications for the management of chronic HBV infections currently available include alpha interferon (IFN-α) and three nucleoside analogs: lamivudine, adefovir and entecavir that inhibit viral nucleocapsid formation and block viral DNA synthesis by premature chain termination [81,82]. The major determinant involved in the selection of drug-resistant mutation is the fitness of the mutants and the replication space available for the spread of mutants. In chronic hepatitis B, the replication space is provided by hepatocyte turnover, which allows the loss of HBV wild-type infected cells and the generation of non-infected hepatocytes that are susceptible to new HBV mutant infections. Long-term therapy of adefovir or entecavir mediates significant reduction in cccDNA, but still fails to eliminate chronic HBV infections [83, 84].

5. Conclusion

Among all forms of viral hepatitis, HBV infection is considered as a major infectious disease due to the broad range of clinical spectrum and the progressive complications displayed by the infected individuals. Avians, rodents and human forms of the virus have been recognized for many years and infections were assumed to be highly host specific. The geographic pattern of HBV genotype distribution is not only influenced by the host and
viral factors but also by socio-economical factors like migration and immigration of people, availability of vaccine and anti-viral therapeutics. The genetic diversity of HBV has been associated with clinical outcome, and response to antiviral therapy. Various forces like natural selection pressure, antiviral drug mediated pressure and error prone high replication rate are the important factors responsible for this genetic diversity. The final outcome from this infectious disease is solely determined by specific interaction between viral components and immunogenetics of the host. Understanding the influence and the role of viral genetic diversity is considered as a prerequisite to better the treatment options.

Author details

Maria Kuttikan Jayalakshmi and Narayanan Kalyanaraman
Department of Immunology, School of Biological Sciences, Madurai Kamaraj University, Madurai, India

Ramasamy Pitchappan
Chettinad University, Kelampakkam, Chennai, India

6. References


