Molecular Biology of the Hepatitis B Virus for Clinicians

Sibnarayan Datta*, Soumya Chatterjee*, Vijay Veer*, Runu Chakravarty**

*Defence Research Laboratory Tezpur, Tezpur, Assam and **ICMR Virus Unit, Kolkata, West Bengal, India

Hepatitis B virus (HBV) infection is one of the major global health problems, especially in economically under-developed or developing countries. HBV infection can lead to a number of clinical outcomes including chronic infection, cirrhosis and liver cancer. It ranks among the top 10 causes of death, being responsible for around 1 million deaths every year. Despite the availability of a highly efficient vaccine and potent antiviral agents, HBV infection still remains a significant clinical problem, particularly in those high endemicity areas where vaccination of large populations has not been possible due to economic reasons.

Although HBV is among the smallest viruses in terms of virion and genome size, it has numerous unique features that make it completely distinct from other DNA viruses. It has a partially double stranded DNA with highly complex genome organization, life cycle and natural history. Remarkably distinct from other DNA viruses, it uses an RNA intermediate called pregenomic RNA (pgRNA) and reverse transcriptase for its genome replication. Genome replication is accomplished by a complex mechanism of *primer shifting* facilitated by direct repeat sequences encoded in the genome. Further, the genome has evolved in such a manner that every single nucleotide of the genome is used for either coding viral proteins or used as regulatory regions or both. Moreover, it utilizes internal *in-frame* translation initiation codons, as well as different reading frames from the same RNA to generate different proteins with diverse functions. HBV also shows considerable genetic variability which has been related with clinical outcomes, replication potential, therapeutic response etc. This review aims at reviewing fundamental events of the viral life cycle including viral replication, transcription and translation, from the molecular standpoint, as well as, highlights the clinical relevance of genetic variability of HBV. (J CLIN EXP HEPATOL 2012;2:353–365)

Globally, hepatitis B virus (HBV) is a major health problem, with almost 2 billion infected subjects, more than 350 million of whom have chronic hepatitis B infection (CHB).¹ CHB has been associated with 100-fold increase in risk for development of hepatocellular cancer (HCC).^{2,3} It is estimated that 15–40% of CHB patients develop severe liver complications such as cirrhosis or HCC in their lifetimes, contributing to more

than 1 million deaths every year. The burden of the disease is increasing as almost 45% of the global population lives in economically developing regions having high prevalence (>8%) of chronic HBV infection, that is, in sub-Saharan Africa and eastern Pacific regions, particularly East Asia. Ironically, even though a highly effective prophylactic vaccine has been available for over three decades, HBV still remains among the 10 leading causes of deaths worldwide.

Although, immediately after its discovery, numerous epidemiological studies strongly established the relation between liver complications and HBV, studies on HBV life cycle were greatly hampered because of its narrow host range, unavailability of HBV supporting cell lines and convenient animal models. Nevertheless, with the advent of modern biomolecular tools, major hurdles have been overcome in the recent decades, which have led to an in depth understanding of HBV virology.⁴ This article aims at reviewing the key facets of HBV molecular biology with special emphasis on virus replication, transcription, translation and genetic variability.

BRIEF HISTORY OF INFECTIOUS HEPATITIS

The earliest description of ‘*epidemic hepatitis*’, which may or may not have been due to HBV, is credited to Hippocrates (ca 450 BC), almost 2500 years ago. In modern times

Keywords: Hepadnavirus, rcDNA, pgRNA, HBsAg, genotype

Received: 17.7.2012; **Accepted:** 12.10.2012; **Available online:** 22.10.2012

Address for correspondence: Dr. Sibnarayan Datta, Scientist ‘C’, Biodefence & Biodiversity Group, Defence Research Laboratory Tezpur (DRDO), Post Bag No. 2, Tezpur Post office, Tezpur 784 001, Assam, India. Tel.: +91 94353 81750; fax: +91 3712 258534

E-mail: sndatta1978@gmail.com

Abbreviations: HBV: hepatitis B virus; HCC: hepatocellular cancer; DR: direct repeat; ORF: open reading frames; RT: reverse transcriptase; TP: terminal protein; LHB: large envelope protein; pHSA: poly-human serum albumin; IL: interleukin; WHV: woodchuck hepatitis virus; rcDNA: relaxed circular DNA; cccDNA: covalently closed circular; BCP: basal core promoter; ER: endoplasmic reticulum; pgRNA: pregenomic RNA; TNF- α : tumor necrosis factor- α ; TGF- α : transforming growth factor- α ; dGMP: deoxyguanosine monophosphate; PC: precore; LEF: liver enriched factors; MHR: major hydrophilic region; EN: enhancer; EBP: enhancer binding protein; CHB: chronic hepatitis B infection; MHBs: middle hepatitis B surface antigen; SHBs: small hepatitis B surface antigen; AUG: translation start codon

<http://dx.doi.org/10.1016/j.jceh.2012.10.003>

Lurman, in a classic epidemiologic study published in 1885, first described 191 cases of 'serum hepatitis' among 1289 shipyard workers in Bremen, Germany, who received small pox vaccine fortified with human lymph.⁵ Large hepatitis epidemics were documented during wars, especially the American Civil War (1861–65), the Franco-Prussian War (1870), World War I (1914–18) and World War II (1939–45).^{6,7} The causative agent for serum hepatitis remained a medical enigma until a geneticist, a hematologist and an airline pilot serendipitously found the 'Australia antigen' (Au) in sera of Australian aborigines in the 1960s and identified it as the hepatitis B virus surface antigen (HBsAg) in 1965.⁸ In 1967, Krugman and colleagues recognized the parenterally transmitted nature of serum hepatitis, which led to the formulation of hygienic measures to prevent serum hepatitis.⁹ The discovery of the Australia antigen was the first step in the development of hepatitis screening assays which dramatically reduced cases of post-transfusion hepatitis and became the basis of a highly effective vaccine that not only controlled hepatitis but also HCC. In 1976, Baruch S. Blumberg was awarded the Nobel Prize in Physiology or Medicine for this landmark discovery of the Au antigen.

BASIC VIROLOGY

Taxonomic Classification

Unique structural features of the HBV set it apart from other families of animal DNA viruses and have led to its classification under a new family—'*Hepadnaviridae*' (hepatotropic DNA viruses). *Hepadnaviridae* family has two genera, the *Orthohepadnaviruses* that infect mammals (human, woodchucks, ground squirrel etc.), and the *Avihepadnaviruses* that infect birds (ducks, wild herons etc.). Human HBV is the prototype member of the *Hepadnaviridae* family.

Viral Morphology

HBV is unique in different aspects of morphology, genetic material, genome organization, genome replication and genetic control.

Unlike other viruses, HBV produces 3 different types of virus-related particles. As visualized by examining preparations from HBV infected persons under the electron microscope, these are—i) spherical, double shelled particles, 42–47 nm in diameter, ii) spherical particles, 20 nm in diameter, and iii) filaments 20 nm in diameter and of variable length (Figure 1). The 42–47 nm double shelled particles, also called *Dane particles* after DS Dane who first described them in 1970, are the actual infectious particles.¹⁰ The number of Dane particles, which determine infectivity of the sample since only they contain the replication competent HBV genome,¹¹ may range from as low as 10^2 particles mL^{-1} of serum in occult and asymptomatic infections to more than 10^8 particles mL^{-1} during the active

Figure 1 Schematic representation of different forms of infectious and non-infectious hepatitis B virus particles. (a) Complete viral particle, (the 42–47 nm Dane particle), (b & c) two species (filamentous and spherical respectively) of non-infectious 20 nm surface antigen particles. 'Pol' indicates HBV polymerase protein. Figure not according to the scale.

replicative phase of infection. Dane particles incorporate 25–27 nm icosahedrally symmetrical nucleocapsids containing the viral nucleic acid, viral polymerase and associated proteins.¹² The 20 nm spherical particles are produced in excess, up to 1000-fold more than the Dane particles, while the 20 nm filamentous particles are produced in lesser amounts.^{4,13} The spherical and filamentous 20 nm particles are mainly composed of the highly immunogenic viral surface glycoprotein, HBsAg, but lack the nucleocapsid with its viral nucleic acid and polymerase activity, so that they are non-infectious.

Viral Nucleic Acid

Another unique feature of HBV is its genome. The HBV DNA is a relaxed, circular, partial double strand, approximately 3.2 kb long. The two strands are asymmetric, a feature exclusive to the *Hepadnaviruses* (Figure 2). The minus strand is complete but contains a 'nick' at a unique site, while the plus strand is incomplete.^{12,14,15} The genome sequence has termini with cohesive ends that match the distinctively located 5' ends of the two strands, and maintain the circular configuration of the DNA.¹⁶ The 5' ends of both the strands incorporate direct repeats (DRs), regions of short repeat sequences, ~11 nucleotides long, which are crucial in priming viral replication.¹⁷ The 5' end of the negative DNA strand encodes the first DR, termed 'DR1', while the positive DNA strand starts with another direct repeat, 'DR2'. The negative strand also has a terminal protein, which is a part of the viral polymerase, covalently linked to its 5' end. On the other hand the plus

Figure 2 Schematic diagram of the genome and translational map of HBV. The innermost red circles depict the rcDNA with the reverse transcriptase/polymerase (Pol) attached to the 5' end of the complete minus strand DNA (solid red circle) and a capped RNA oligomer (wavy red line) attached to the 5' end of the incomplete plus strand DNA (dotted red circle). The positions of the direct repeats (DR1 and DR2) and enhancers (EN1 and EN2) are also indicated. The green line indicates the viral genome positions in nucleotides (approximate). The four protein-coding regions are shown between the green and red circles by colored semi-circular arrows. They include the precore (PC) and core genes (violet), the polymerase gene (blue), the X gene (aqua) and the envelope genes preS1, preS2, and S (orange). Genome positions may change depending upon the genotype of the HBV genome (modified from Nassal 2008). The outermost semi-circular lines with arrowheads represent the 4 RNAs (genomic and subgenomic) corresponding to the ORFs. The arrowheads indicate the positions of different initiation codons within each ORF.

strand has a capped oligoribonucleotide attached at its 5' end (Figure 2).

Viral Genomic Organization

The HBV DNA has a highly compact and highly evolved organization of the coding and regulatory regions, where every nucleotide within the genome is utilized in one or the other of the coding regions or open reading frames (ORFs).^{18,19} HBV genome codes four highly overlapping ORFs (Figure 2). Viral polymerase, the central enzyme in genome replication, is encoded by the P (Pol) ORF. It has DNA polymerase (DNA Pol), reverse transcriptase (RT) and RNase H activities and also acts as the terminal protein (TP).²⁰ The C (core) gene encodes the structural protein of the nucleocapsids as well as the 'e' antigen (HBeAg). The S (surface) region encodes 3 different envelope glycoproteins. The X region encodes the multifunctional X protein. The genome also contains regulatory elements (promoters, enhancers) and also encodes signals for polyadenylation and encapsidation.⁴

Coding of different virion proteins using the same ORF is done by exploiting more than one in-frame initiation codon. Due to this fact, three different surface molecules, with variable N-terminals, but a common C-terminal end, are synthesized from the preS/S ORF. The largest of these, the preS1 protein (or large HBsAg or LHBs) is the product of initiation at the first initiation codon of the ORF, while initiation at the second initiation codon produces preS2 middle hepatitis B surface antigen (or middle HBsAg or MHBs). Initiation at the third start codon produces the classical HBsAg small hepatitis B surface antigen (or small HBsAg or SHBs) that contains only the S domain, and is commonly referred to as the surface antigen (HBsAg or Au antigen). The core protein gene also has at least two in-frame start codons, with the HBcAg being produced from the internal translation start codon (AUG), while initiation at the upstream AUG gives rise to a C related soluble protein, which is secreted into the circulation as hepatitis B e antigen (HBeAg). Although less characterized, usage of two different in-frame AUGs in production of

N-terminal variable related proteins has also been reported for the X ORF.²¹

HEPADNAVIRAL LIFE CYCLE

Hepatitis B virus (HBV) has evolved unique mechanisms to actively replicate in vulnerable cells without killing the infected cell directly. One of these unique mechanisms is that the HBV genome replicates via reverse transcription of an RNA intermediate, instead of DNA replication through semi-conservative replication.^{22–24} Although the mechanism of genome replication has now been meticulously studied, our knowledge of virion entry, uncoating, and nuclear transport of viral genome is still inadequate. Different steps of the viral life cycle and genome replication are described in the following sections (Figure 3).

Infection, Attachment and Viral Cell Entry

Classically, HBV is believed to be highly hepatotropic and hepatocytes are considered as the primary site of viral infection. This tissue specificity is thought to be due to one or more poorly characterized hepatocyte-specific HBV receptors and/or co-receptors. Pseudotyping experiments and other studies indicate that the preS1 protein (specifically, amino acids 21–47) participates in cellular receptor binding.^{25–27} A variety of human cellular proteins have been shown to bind HBV envelope glycoproteins. Among them serum apolipoprotein H (apo H), poly-human serum albumin (pHSA), carboxypeptidase D, fibronectin and interleukin-6 (IL-6) have been suggested as putative candidate receptors for HBV.^{28–31} After binding, entry of the virus into the cell occurs by a host-virus membrane fusion

event. It is believed that entry occurs prior to endocytosis through a pH independent mechanism.

Very interestingly, recent experiments on woodchuck hepatitis virus (WHV) models have clearly established that the woodchuck hepatitis virus, an animal model of HBV, has almost 1000 times higher affinity toward infecting lymphoid cells as compared to hepatocytes.³² These observations strongly challenge classical thought and signify that WHV might actually be a lymphotropic virus. Lymphotropism has been documented in the case of HBV also, but the mechanism, life cycle and clinical significance of the virus in lymphocytes are yet to be elucidated.

Genome Repair and Gene Expression

Very little is known about nuclear transport of the core particles after endocytosis. It is believed that nuclear transportation is directed by a nuclear localization signal in the C terminus of the core particle. However, cytoplasmic localization of the core particle implies that it is held in the cytosol while the viral DNA along with the polymerase is conveyed into the nucleus in a phosphorylation dependent event. Subsequently, relaxed circular DNA (rcDNA) is repaired by completion of the plus strand, removal of 5' terminal structures (terminal protein in the minus strand, oligoribonucleotide primer in the plus strand), and ligation of the respective strands, to generate a supercoiled covalently closed circular (cccDNA).^{4,13} These repair reactions are primarily executed by the host cellular repair mechanisms and are independent of viral polymerase activity.³³ Formation of cccDNA is the hallmark of establishment of infection and it serves as the template for transcription of viral genomic and subgenomic RNAs. Successive steps of the HBV life cycle are shown in Figure 3.

Figure 3 Schematic diagram showing life cycle of the HBV.

Hepadnaviruses use the negative strand nucleotide sequence of cccDNA as template for transcription of genomic and subgenomic RNAs, all being 5' capped and 3' polyadenylated at a common polyadenylation signal located in the core ORF (Figure 2). The three subgenomic RNA transcripts are the 2.4 kb, 2.1 kb and 0.7 kb RNAs, which are translated, respectively, into envelope preS1 protein, preS2 and HBsAg proteins and the X protein.^{34,35} The genomic RNA, translates into the precore, core and polymerase proteins (Figure 2). It is also popularly known as pregenomic RNA (pgRNA), since it is also the template for synthesis of viral DNA negative strand through reverse transcription. Even though HBV mRNAs are not known to encode introns or other non-coding regions, spliced viral transcripts have been detected *in vivo*.^{36,37} Transcription of the four HBV ORFs is tightly regulated by enhancer II (EN2)/basal core promoter (BCP), large surface antigen, major surface antigen, X gene promoters and enhancer I (EN1).⁴ With a very compact, highly overlapping genome, *Hepadnaviruses* fully utilize each nucleotide in their genome. Not only do certain stretches of the genome encode two or more different proteins by the frame shift mechanism, the same sequences also act as regulatory regions.

Viral Genes and the Functions of the Proteins

Apart from using different coding frames, HBV also utilizes multiple initiation codons to generate structurally related, functionally diverse proteins from the same reading frames. In the preC/C ORF, translation initiating from the 5' initiation codon translates into the precore protein while translation initiating at the 3' initiation codon translate into the core protein. Due to the presence of a signal peptidase recognition motif, the preC polypeptide is cleaved in the endoplasmic reticulum (ER) and secreted as a 17 kD HBeAg, which acts as an immunomodulatory molecule. This protein has been shown to induce immune tolerance at the T-cell level. HBeAg shares the same T-cell epitopes as the core protein, hence tolerance induced by HBeAg also induces tolerance to the core protein which precludes the elimination of HBV infected cells by T-cell mediated pathways.³⁸

The core ORF encodes the ~21 kD viral core or capsid protein (HBcAg). Its N-terminal is essential for self-assembly (assembly domain), while the C-terminal is important for capsid formation, interaction with pgRNA/reverse transcriptase complex and the packaging of the complex inside capsids. Capsid formation is brought about by the arrangement of homodimers of the core protein and their stabilization through cysteine residues, finally resulting in the formation of icosahedral subviral capsids.³⁹ Two types of capsids have been observed, i) icosahedral capsids of T = 3 symmetry (30 nm diameter, 180 core particles), and ii) icosahedral capsids of T = 4 symmetry (34 nm diameter, 240 core particles), the latter being found chiefly in infectious virions.^{40–42}

Interestingly, almost 80% of the viral genome is used to encode the viral polymerase polypeptide (also referred to as Pol), which overlaps enormously with all the other ORFs. The Pol is translated from pgRNA using an internal initiation codon. This large (~94 kD) multifunctional protein consists of at least four domains and has three discrete enzymatic functions. The N-terminal domain termed *primase* (or terminal protein, TP domain) acts in priming negative DNA strand synthesis and also plays a critical role in the packaging of pgRNA while the second or '*spacer*' domain acts as a *spacer* and might have some other important functions which are yet unknown. The third or *polymerase* domain has RNA- and DNA-dependent polymerase activity, crucial for genome replication, and the fourth domain having *RNase H* activity, is essentially required during genome replication.¹³

The HBV envelope proteins are expressed from the preS/S ORF which has at least three distinct in-frame translation initiation sites. Translation of the largest subgenomic mRNA (2.4 kb) yields the large envelope protein (L) while translation of the shorter transcript (2.1 kb) gives rise to the middle (M) and small surface proteins (S or HBsAg), depending on which of the two in-frame initiation codons is activated. All three envelope proteins are glycosylated type II transmembrane proteins and are stabilized by disulfide bridges. The N-terminal domain unique to the L protein is called the preS1 domain, while N-terminal domain of M is called the preS2 domain. The preS1 domain is the substrate for viral receptor attachment and epitopes displayed outside can elicit virus-neutralizing antibodies.⁴³ PreS1 epitopes displayed on the cytosolic side of the viral membrane acts as ligands for interaction with core particles during the assembly of the viral envelope. The function of the M protein is not well understood. On the other hand, the S or HBsAg protein contains the major B cell epitope, the "a" determinant (residues 121–149 aa) which maps to an external, hydrophilic region containing highly conserved cysteine residues. HBsAg appears as the first marker during infection and may persist in patients who progress to CHB.

The fourth, or the X, ORF is exclusive to *Orthohepadnavirus*, the genus infecting mammals and is absent in the genus *Avihepadnavirus*, infecting birds. The functions of the HBx protein during natural viral infection are still unclear. It is thought to be essential for successful establishment of infection and replication in animals but not in transfected tissue culture cells.^{44,45} Antibodies are detectable in sera of naturally infected individuals, indicating that it is indeed expressed *in vivo* during infection. *In vitro*, HBx exhibits numerous activities, from transactivation of host and viral genes, to stabilization of viral RNAs and is believed to function through protein-protein interactions.^{45,46} HBx is capable of transactivation of viral and a range of cellular genes, stimulation or inhibition of various signal transduction pathways, binding to a variety of cellular protein targets, modulation of protein degradation, DNA

repair, apoptosis, inhibition of tumor necrosis factor- α (TNF- α), and transforming growth factor- α (TGF- α) induced apoptosis, to name a few. These observations explain the contribution of HBx toward the development of HCC through alteration of different signaling pathways by different mechanisms.^{45–48} In addition, the C-terminal domain of the HBx has been shown to complex with p53 in the cytoplasm, partially preventing its nuclear entry and ability to induce apoptosis, contributing to early stages of HCC development.⁴⁹

HBV envelope glycoproteins are also associated with transformation and HBV induced diseases. Specifically, accumulation of large envelope protein (LHBs) within the ER of hepatocytes has been shown to be associated with predisposition to transformation.⁵⁰ Furthermore, studies showed that HBV DNA integrated in the host genome can express an MHBs protein, which is truncated at the carboxy terminal end, termed as MHbSt, capable of transactivation of cellular gene expression.⁵¹ However, the exact role of these variants in hepatocarcinogenesis remains to be elucidated.

Encapsidation of Pregenomic RNA

As stated earlier, pgRNA serves as mRNA for the viral pre-core, core and Pol proteins. Pol initiation starts from an internal AUG codon and the reading frame is +1 relative to the C frame.⁵² Interestingly, due to some yet unknown reasons, despite being translated from the same pgRNA, reading of the polymerase ORF is very inefficient as compared to the reading of the core ORF. It has been reported that roughly 200–300 core proteins are translated before the pgRNA is available for production of only a couple of Pol proteins.^{40,53} Whatever be the mechanism, this preferential reading and translation of the core protein ensures a correct ratio of core and Pol proteins and prevents the accumulation of viral nucleic acids to cytotoxic levels. Synthesis of polymerase signals translation, termination and initiation of viral packaging. The viral polymerase binds to the *cis*-acting packaging signal (a stem loop structure termed as *epsilon*, 'e') present at the 5' end of pgRNA, prompting packing of viral RNA and polymerase into core particles.^{54–57} Due to its unique spatial conformation, ϵ stem loop also induces structural changes in both the pgRNA and polymerase, to initiate its own reverse transcription.¹³ Although the 'e' element is present in both the ends of the pgRNA due to terminal redundancy, for some reason, only the 5' copy is essential for signaling.⁵⁸ Interestingly, the exclusive requirement of 5' ϵ for signaling encapsidation and replication ensures that only the pgRNA is packed and transformed to rcDNA inside viral particles, since other subgenomic RNAs lack the 5' copy of the ϵ signal. RNA packaging also depends on host factors, in particular the *hsp90* molecular chaperone complex.^{4,13}

Replication of the Genome

The method of converting the single stranded pgRNA to partially double stranded rcDNA is highly complex and again unique for *Hepadnaviruses* (shown in Figure 4a–f). At the initial stage of genome replication, the ϵ epsilon stem loop employs an N-terminal hydroxyl group of a tyrosine residue of the TP domain of polymerase protein as a starting point to add nucleotides.^{20,59–61} This hydroxyl group serves as a substrate for formation of a phosphodiester linkage with a dGMP (deoxyguanosine monophosphate) and further addition of 4 nucleotides using bulge region of epsilon stem loop RNA sequence as a template.^{58,62,63} This nascent DNA strand is then transferred, by a mechanism yet to be understood, to the complementary sequences (*first primer shift*), also present at the 3' end of the pgRNA.²² Subsequently, reverse transcription of the pgRNA takes place and minus DNA strand is synthesized in 5' to 3' direction and, simultaneously, pgRNA template is degraded by the RNase H activity of the Pol protein.²² The use of one of the hydroxyl groups of the TP domain of Pol results in covalent linkage between the 5' end of minus strand DNA and the polymerase protein.⁶⁴ On the other hand, despite being mostly degraded by the RNase H activity, a 15–18 nt long stretch of 5' capped end of pgRNA remains undegraded, which also includes the DR1 sequence. This oligoribonucleotide acts as an RNA primer for plus strand synthesis by reallocating base pairs to the complementary 5' DR2 region on minus DNA strand (*second primer shift*); the plus strand then elongates toward the 5' end of minus strand DNA.^{17,65,66} This combined apparatus of reverse transcription and RNase H activity ensures that the minus strand is synthesized to its completion along with degradation of the pgRNA template and subsequent generation of the RNA primer for initiation of the plus strand.^{65,66} With ongoing elongation of the plus strand, at a point when the minus strand DNA template is exhausted, an intramolecular strand transfer occurs (*template shift*). This is made possible by the redundant sequences present in the negative strand DNA template. Annealing of the nascent plus strand DNA by complementary base pairing with the redundant portion of the 3' end of the minus strand DNA circularizes the genome and also allows the plus strand to access rest of the minus DNA as template for elongation.^{67,68} It is believed that post-translational modifications of the core particles occur simultaneously with genome maturation and replication, which also signals capsid envelopment and budding into ER. Once inside the core particle, the viral replication apparatus has limited access to the cytoplasmic dNTP pool and elongation of the plus strand is arrested or terminated prematurely due to the shortage of dNTPs. The final result of the genome replication is a rcDNA genome with a partially double stranded DNA, a hallmark of the *Hepadnavirus*

Figure 4 Consecutive steps of HBV genome replication. (a) Translation of pgRNA and initiation of minus strand reverse transcription using the TP domain of Pol and bulge region of 'ε' as template. (b) First primer shift-translocation of nascent minus strand to 3' end of pgRNA and base pairing with DR2. (c) Completion of minus strand using pgRNA as template and simultaneous degradation of pgRNA. (d) Second primer shift and template shift—reallocation of RNA primer to the 5' end of the minus strand (DR2), followed by circularization of the genome due to redundant sequences present in the minus strand DNA. (e) Formation of relaxed circular DNA (rcDNA)—incomplete elongation of the plus strand DNA using minus strand DNA as template. (f) Infection and repair of rcDNA to covalently closed circular DNA.

genome.^{12,14,15,69} The encapsidated rcDNA thus formed may either enter the nucleus for maintenance of the cccDNA pool, or follow the secretion pathway to be exuded as an infectious virion.

Rarely, RNA primer translocation to the complementary 5' DR2 region on minus DNA strand (*second primer shift*) does not occur, resulting in a situation termed as '*in situ priming*' that generates a double stranded linear DNA (dslDNA) genome instead of the rcDNA genome.⁷⁰ Although the significance of dslDNA in natural infection is not known, it has been shown to be infectious, capable of cccDNA formation, viral DNA synthesis, and are also the preferred form for integration into host genome.^{71,72}

Amplification of Covalently Closed Circular DNA Pool in Nucleus and Production of Virus Particles

One of the most fascinating characteristics of the *Hepadnavirus* life cycle is the regulation of cccDNA concentration,

which is the basis of persistence and maintenance of infection in the susceptible host cells. It has been reported that cccDNA accumulates to toxic levels in the absence of envelope proteins, indicating a feedback loop type regulation of cccDNA concentration.^{73,74} It is thought that cccDNA amplification preferentially occurs during the early phase of an infection, when the concentration of the envelope protein is low. It is estimated that 5–30 copies of cccDNA are present per nucleus *in vivo*.^{75,76}

With ongoing infection, surface gene transcripts direct synthesis of different types of surface proteins (SHBs, MHBs, and LHBs). These proteins have transmembrane spanning domains which help them in budding into endoplasmic reticulum (ER) followed by secretion as 20–22 nm non-infectious particles. Simultaneously, through interaction with envelope polypeptides residing in ER membranes, nucleocapsids are also led into the secretory pathway. Interaction of core particles with N-terminal domain of the large envelope polypeptides results in budding

of the core particles into the ER membrane. Subsequently, enveloped 42 nm particles containing the viral genome are transported through the ER into the Golgi complex.⁷⁷ Glycosylation of the envelope proteins occurs during the assembly process and mature virions are released to infect the neighboring cells or to reach remote locations through the bloodstream.

GENETIC VARIABILITY OF HEPATITIS B VIRUS

As it is for other viruses, genetic variability and genome evolution is an important attribute of the HBV life cycle. However, unlike other DNA viruses, HBV has certain distinct features which set it apart. Notably, the HBV polymerase lacks proofreading function, and uses an RNA intermediate (pgRNA) during its replication. These two features enhance the probability of random errors during genomic replication, leading to evolution of the genome.^{18,19,78} In general, nucleotide substitution rate for HBV is estimated to be $1.4 - 5.0 \times 10^{-5}$ per site per year, almost 10 times higher than for other DNA viruses and similar to certain RNA viruses.⁷⁹ On the other hand, due to highly overlapping ORFs in the HBV genome, an advantageous variant in one ORF may produce non-advantageous or even lethal effects in the another. Hence, the evolution of HBV is constrained and the rate of synonymous (silent) substitutions is higher than the rate of non-synonymous substitutions in all four ORFs.^{80,81} However, the rate of evolution and emergence of variants may greatly differ in different clinical settings, depending upon the pressure exerted by the host immune system, vaccine immunophylaxis or antiviral drugs. It has been reported that mutation rate is almost 100-fold higher in immune-suppressed patients as compared to subjects having silent or occult HBV infection.^{82,83} Nevertheless as a result of complex host-virus interactions, mutations emerging under different types of selective pressure are either fixed in the genome if advantageous or eliminated during subsequent cycles if disadvantageous.

HBV genetic variability can also be grouped into 2 types namely genotype/subgenotype variability and random mutations.

Hepatitis B virus Genotype/Subgenotype

HBV genotypes or subgenotypes signify stable variants that have arisen due to selection and fixation of random changes over long periods of time in particular human populations. These strains circulate in a stable manner and are broadly restricted to the host population in which they have evolved, thus having distinct geo-ethnic distribution patterns. Classically, HBV strains were discriminated as *serotypes*, based on the main antigenic determinant 'a' and two mutually exclusive determinants 'd' and 'y' in the surface protein.^{18,84} Two additional determinants, 'w'

and 'r', were later added and HBV strain was categorized as serotype 'adw', 'adr', 'ayw' or 'ayr'. Further determinants were subsequently described leading to the current classification with nine serotypes namely *ayw1*, *ayw2*, *ayw3*, *ayw4*, *ayr*, *adw2*, *adw4*, *adrq+* and *adrq*.¹⁸

Once detailed phylogenetic analysis became possible following the availability of complete HBV genome sequences, Okamoto et al classified the HBV genomes into 4 distinct genotypes, A to D, based on more than 8% inter-group divergence between the genotypes.⁸⁵ Increasing information on HBV genome sequences from different parts of the globe and their extensive phylogenetic analyses led to the identification of other genotypes, and resulted in further classification into subgenotypes, based on more than 4% but less than 8% intragenotype divergence in nucleotide sequences.^{84,86,87} At present, 7 subgenotypes of genotype A, 9 of genotype B, 12 of genotype C, 7 of genotype D and 4 of genotype F have been described.⁸⁸⁻⁹¹ Recently two more genotypes, genotype I and genotype J have been proposed for two recombinant strains of HBV, reported from Vietnam and Japan respectively.^{92,93} Different genotypes are believed to have evolved from a common ancestor, independently and in parallel. It has also been suggested that extensive recombination events have played important roles in the evolution of different genotypes of HBV.⁹⁴ Presently, with the availability of powerful computational programs, comparative phylogenetic analysis of HBV sequences have become pervasive in studies related to anthropological migration patterns, routes of infection and transmission patterns.^{18,87}

A fascinating development has been the realization that there are significant differences in clinical behavior patterns among different HBV genotypes and subgenotypes. The initial proof of clinical differences between different HBV genotypes was provided by Tabor and co-workers, who clearly showed different course of disease for HBV serotypes *ayw*, *adr* and *adw* which were actually encoded by different genotypes.⁹⁵ More interestingly, the endemicity of HBV infection and the prevalence of liver disease also vary geographically. In East Asian countries, where genotypes B, C are prevalent and in sub-Saharan countries, where genotype A is prevalent, HBV are found to be hyperendemic and the incidence of HCC is also high.⁸⁷ The mode of HBV transmission also varies according to HBV genotypes. Horizontal transmission is common in sub-Saharan Africa, the Mediterranean, the Middle East and India where genotypes A and D are prevalent, whereas vertical transmission is common in East Asian-Pacific regions where genotypes B and C are prevalent.^{96,97} Comparative analysis of HBV genotype distribution among different patient groups in a particular population, where more than one HBV genotypes are prevalent, have clearly shown that certain HBV genotypes are significantly associated with emergence of persistent HBV variants, clinical severity, recovery patterns or response to antiviral treatment.^{19,86,87,98,99}

The most comprehensively analyzed mutations of HBV are those appearing in BCP and PC regions. Among these, double mutations in BCP region (mutations 1762T/1764A), which regulates HBeAg synthesis at transcriptional level and a single mutation in the PC region (1896A), which regulates HBeAg synthesis at the translational level, are the most studied.¹⁰⁴ These mutations also influence stability of the secondary structure of pgRNA and thus its replication competence.¹⁰⁴ The 1762T/1764A double mutation (Adenine to Thymine at position 1762 and Guanine to Adenine at position 1764) results in significant transcriptional reduction of precore and core mRNA. These double mutations are often detectable in patients with chronic hepatitis, fulminant hepatitis or hepatocellular carcinoma but less frequently noted among asymptomatic carriers.¹⁰⁵ 1762T/1764A lies within CCAAT/Enhancer binding protein (C/EBP) binding region and decreases binding of liver enriched factors (LEF), lowering HBeAg production. On the other hand, a Guanine to Adenine substitution at position 1896 (denoted as 1896A), located in the epsilon (ϵ) structure of the precore gene, converts the 28th codon of HBeAg from Tryptophan to translational stop codon (TGG to TAG). The ϵ structure is a highly conserved stem loop RNA structure and mutations at a certain site also influence the variability of another site that base pairs with it and vice-versa. It has been shown that 1896G base pairs with 1858C at the base of the stem loop (Figure 5). Interestingly, nucleotide at position 1858 is HBV genotype specific. Hence emergence of 1896A is common in HBV genotypes having 1858T and less frequent in HBV genotypes having 1858C. These mutations are frequently detected among individuals with HBeAg negative active infection.

Among the HBV ORFs, the surface ORF has been found to be extremely heterogeneous.^{50,106} Numerous point mutations, deletions and recombinations have been identified in the preS region. Apart from the preS variants, a number of mutations have also been detected within the “a” determinant of the major hydrophilic region (MHR) of the surface antigen, against which natural or vaccine induced neutralizing antibodies are generated. Mutations within this epitope, which help the virus to escape immune surveillance, are selected. Two most widely studied immune escape mutations are G145R (glycine to arginine at residue 145 of HBsAg) and D144A (aspartate to alanine at residue 144).^{107,108} Interestingly, despite the high efficacy of HBV vaccine in prevention of HBV infection, breakthrough infections due to vaccine escape mutations have been reported in vaccinated individuals, which accentuate the importance of these escape mutations.^{109,110}

With the increasing use of nucleoside and nucleotide analogs as HBV antiviral agents, reports of antiviral resistance mutations in the HBV Pol gene have also increased in the recent years (Figure 6). The most studied mutation in this category is the lamivudine resistance primary mutation rtM204V/I in domain C, (changing the conserved YMDD motif to YVDD or YIDD), and is usually accompanied by another compensatory or secondary mutation upstream of the YMDD motif at rtL180M and/or rtV173L in the domain B of HBV polymerase. This secondary mutation is reported to enhance the replicative efficiency of the resistant mutant.¹¹¹ On the other hand, resistance to adefovir dipivoxil is associated with mutations rtN236T

Figure 6 Schematic diagram showing the functional domains of the HBV polymerase (Pol) protein, and the active domains within the polymerase/reverse transcriptase (RT) domain. 'TP' indicate terminal protein. Position of different antiviral resistance mutations within the polymerase/RT domain is also shown below [adapted and modified from Locarnini and Yuen, 2010].

in the domain D and rtA181T/V in domain B.^{112,113} Entecavir resistance mutations rtT184S/A/I/G/C/M, rtS202I/C/G, and rtM250I/V have been mapped to domains B, C, and D respectively. More comprehensive reviews on drug resistant and immune escape mutants of HBV have been published elsewhere.^{114,115}

Finally, nearly all deletions or insertions in the basal core promoter shift the X gene frame and lead to the production of truncated X proteins. These shortened X proteins lack the domain in the C terminus (amino acids 130–140) that is required for the transactivation activity of HBx antigen. Other clinically important HBx variants/mutations have been discussed elsewhere.¹¹⁶

CONCLUSION

During the past decades, rapid advances in animal models, molecular biology and immunology have resolved many issues in understanding hepatitis B virus life cycle and pathogenesis. We now know a great deal about the complex interactions between the host immune system and the virus, although information is not complete and some important facets of the HBV life cycle still

remain a riddle. A major lacuna is the failure in identification of cellular receptors that mediate infection and tissue tropism. Although template switches during the hepadnaviral DNA synthesis have been recognized, exact mechanisms and stimuli for such events are still poorly understood. The precise mechanisms of cccDNA formation and persistence in the nucleus are not properly understood. Information about the specific role of viral proteins, especially HBx, in disease pathogenesis and carcinogenesis, still remains inconclusive, despite the fact that numerous studies have suggested numerous pathways and interactions with cellular proteins. Although there are clues about the clinical significance of one, two or perhaps three related mutations, direct correlation between a combination of unrelated viral mutations and specific clinical outcomes cannot be made.

Recent studies have added yet another dimension to this complexity, namely limited understanding of differences between viral genotypes and subgenotypes. Epidemiological studies have now clearly established that there are genotype-specific differences in HBV transmission patterns and outcomes of HBV infection such as recovery from acute infection, chronicity and even response to

therapy. This raises serious concerns that the present understanding of HBV replication, tropism and pathogenicity may not be universally applicable to all genotypes of HBV. Genotype-specific differences in HBV life cycle and pathogenicity need to be recognized.

In conclusion, further in depth studies are of utmost importance to resolve these issues. These may help in developing better prophylactic approaches, novel biomarkers as well as in identifying new therapeutic targets and genotype-specific interventional approaches that may help millions of patients infected currently with HBV or likely to get infected in the future.

CONFLICTS OF INTEREST

All authors have none to declare.

ACKNOWLEDGMENTS

The authors thank Defence Research and Development Organization (DRDO), Ministry of Defence, Govt. of India and Indian Council of Medical Research (ICMR), Ministry of Health and Family Welfare, Govt. of India for financial support.

REFERENCES

- European Association for the Study of the Liver. EASL Clinical Practice Guidelines: management of chronic hepatitis B. *J Hepatol*. 2012;57:167–185.
- Beasley RP, Hwang LY, Lin CC, Chien CS. Hepatocellular carcinoma and hepatitis B virus. A prospective study of 22,707 men in Taiwan. *Lancet*. 1981;2:1129–1133.
- Shepard CW, Simard EP, Finelli L, Fiore AE, Bell BP. Hepatitis B virus infection: epidemiology and vaccination. *Epidemiol Rev*. 2006;28:112–125.
- Seeger C, Mason WS. Hepatitis B virus biology. *Microbiol Mol Biol Rev*. 2000;64:51–68.
- Lurman A. Eine icterus epidemic. *Berl Klin Wochenschr*. 1885;22:20–23 [in German].
- Sherlock S, Dooley J. Viral hepatitis. In: Sherlock S, Dooley J, eds. *Disease of the Liver and Biliary System*. 10th ed. Milan: Blackwell Science Ltd; 1997:265–302.
- Zuckerman AJ. The chronicle of viral hepatitis. *Bull Hyg Trop Dis*. 1997;54:113.
- Blumberg BS, Alter HJ, Visnich SA. A new antigen in leukemia sera. *JAMA*. 1965;191:541–546.
- Krugman S, Giles JP, Hammond J. Infectious hepatitis: evidence for two distinctive clinical, epidemiological, and immunological types of infection. *JAMA*. 1967;200:365–373.
- Dane DS, Cameron CH, Briggs M. Virus-like particles in serum of patients with Australia-antigen-associated hepatitis. *Lancet*. 1970;1:695–698.
- Robinson WS, Clayton DA, Greenman RL. DNA of a human hepatitis B virus candidate. *J Virol*. 1974;14:384–391.
- Kaplan PM, Greenman RL, Gerin JL, Purcell RH, Robinson WS. DNA polymerase associated with human hepatitis B antigen. *J Virol*. 1973;12:995–1005.
- Ganem D, Schneider RJ. Hepadnaviridae: the viruses and their replication. In: Knipe DM, Howley PM, eds. *Fields Virology*. 4th ed. Philadelphia: Lippincott Williams & Wilkins; 2001:2923–2969.
- Summers J, O'Connell A, Millman I. Genome of hepatitis B virus: restriction enzyme cleavage and structure of DNA extracted from Dane particles. *Proc Natl Acad Sci U S A*. 1975;72:4597–4601.
- Landers TA, Greenberg HB, Robinson WS. Structure of hepatitis B Dane particle DNA and nature of the endogenous DNA polymerase reaction. *J Virol*. 1977;23:368–376.
- Delius H, Gough NM, Cameron CH, Murray K. Structure of the B virus genome. *J Virol*. 1983;47:337–343.
- Haines KM, Loeb DD. The sequence of the RNA primer and the DNA template influence the initiation of plus-strand DNA synthesis in hepatitis B virus. *J Mol Biol*. 2007;370:471–480.
- Kidd-Ljunggren K, Miyakawa Y, Kidd AH. Genetic variability in hepatitis B viruses. *J Gen Virol*. 2002;83:1267–1280.
- Kay A, Zoulim F. Hepatitis B virus genetic variability and evolution. *Virus Res*. 2007;127:164–176.
- Bartenschlager R, Schaller H. The amino-terminal domain of the hepadnaviral P-gene encodes the terminal protein (genome-linked protein) believed to prime reverse transcription. *EMBO J*. 1988;7:4185–4192.
- Kwee L, Lucito R, Aufiero B, Schneider RJ. Alternate translation initiation on hepatitis B virus X mRNA produces multiple polypeptides that differentially transactivate class II and III promoters. *J Virol*. 1992;66:4382–4389.
- Summers J, Mason WS. Replication of the genome of a hepatitis B-like virus by reverse transcription of an RNA intermediate. *Cell*. 1982;29:403–415.
- Ganem D, Pollack JR, Tavis J. Hepatitis B virus reverse transcriptase and its many roles in hepadnaviral genomic replication. *Infect Agents Dis*. 1994;3:85–93.
- Seeger C, Mason WS. Replication of the hepatitis virus genome. In: DePamphilis ML, ed. *DNA Replication in Eukaryotic Cells*. New York: ColdSpringHarbor Laboratory Press; 1996:815–831.
- Pontisso P, Morsica G, Ruvoletto MG, et al. Hepatitis B virus binds to peripheral blood mononuclear cells via the pre S1 protein. *J Hepatol*. 1991;12:203–206.
- Petit MA, Capel F, Dubanchet S, Habit H. PreS1-specific binding proteins as potential receptors for hepatitis B virus in human hepatocytes. *Virology*. 1992;187:211–222.
- Dash S, Rao KVS, Panda SK. Receptor for pre-S1 (21–47) component of hepatitis B virus on the liver cell: role in virus cell interaction. *J Med Virol*. 1992;37:116–121.
- Mehdi H, Kaplan MJ, Anlar FY, et al. Hepatitis B virus surface antigen binds to apolipoprotein H. *J Virol*. 1994;68:2415–2424.
- Pontisso P, Petit MA, Bankowski MJ, Peeples ME. Human liver plasma membranes contain receptors for the hepatitis B virus preS1 region and, via polymerized human serum albumin, for the preS2 region. *J Virol*. 1989;63:1981–1988.
- Kuroki K, Eng F, Ishikawa T, Turck C, Harada F, Ganem D. gp180, a host cell glycoprotein that binds duck hepatitis B virus particles is encoded by a member of the carboxypeptidase gene family. *J Biol Chem*. 1995;270:15022–15028.
- Neurath AR, Strick N, Sproul P. Search for hepatitis B virus cell receptors reveals binding sites for interleukin 6 on the virus envelope protein. *J Exp Med*. 1992;175:461–469.
- Jin YM, Churchill ND, Michalak TI. Protease-activated lymphoid cell and hepatocyte recognition site in the preS1 domain of the large woodchuck hepatitis virus envelope protein. *J Gen Virol*. 1996;77:1837–1846.
- Kock J, Schlicht HJ. Analysis of the earliest steps of hepadnavirus replication: genome repair after infectious entry into hepatocytes does not depend on viral polymerase activity. *J Virol*. 1993;67:4867–4874.
- Cattaneo R, Will H, Schaller H. Hepatitis B virus transcription in the infected liver. *EMBO J*. 1984;3:2191–2196.

35. Enders GH, Ganem D, Varmus H. Mapping the major transcripts of ground squirrel hepatitis virus: the presumptive template for reverse transcriptase is terminally redundant. *Cell*. 1985;42(1):297–308.
36. Wu HL, Chen PJ, Tu SJ, Lin MH, Lai MY, Chen DS. Characterization and genetic analysis of alternatively spliced transcripts of hepatitis B virus in infected human liver tissues and transfected HepG2 Cells. *J Virol*. 1991;65:1680–1686.
37. Soussan P, Tuveri R, Nalpas B, et al. The expression of hepatitis B spliced protein (HBSP) encoded by a spliced hepatitis B virus RNA is associated with viral replication and liver fibrosis. *J Hepatol*. 2003;38:343–348.
38. Milich DR, Chen MK, Hughes JL, Jones JE. The secreted hepatitis B precore antigen can modulate the immune response to the nucleocapsid: a mechanism for persistence. *J Immunol*. 1998;160:2013–2021.
39. Chang C, Zhou S, Ganem D, Standring DN. Phenotypic mixing between different hepadnavirus nucleocapsid proteins reveals C protein dimerization to be cis preferential. *J Virol*. 1994;68:5225–5231.
40. Crowther RA, Kiselev NA, Böttcher B, et al. Three-dimensional structure of hepatitis B virus core particles determined by electron cryomicroscopy. *Cell*. 1994;77:943–950.
41. Zlotnick A, Cheng N, Conway JF, et al. Dimorphism of hepatitis B virus capsids is strongly influenced by the C-terminus of the capsid protein. *Biochem*. 1996;35:7412–7421.
42. Wynne SA, Crowther RA, Leslie AG. The crystal structure of the human hepatitis B virus capsid. *Mol Cell*. 1999;3:771–780.
43. Heermann KH, Kruse F, Seifer M, Gerlich WH. Immunogenicity of the gene S and pre-S domains in hepatitis B virions and HBsAg filaments. *Intervirology*. 1987;28:14–25.
44. Chen DS. From hepatitis to hepatoma: lessons from type B viral hepatitis. *Science*. 1993;262:369–370.
45. Bouchard MJ, Schneider RJ. The enigmatic X gene of hepatitis B virus. *J Virol*. 2004;78:12725–12734.
46. Tang H, Oishi N, Kaneko S, Murakami S. Molecular functions and biological roles of hepatitis B virus X protein. *Cancer Sci*. 2006;97:977–983.
47. Murakami S. Hepatitis B virus X protein: structure, function and biology. *Intervirology*. 1999;42:81–99.
48. Pang R, Tse E, Poon RTP. Molecular pathways in hepatocellular carcinoma. *Cancer Lett*. 2006;240:157–169.
49. Elmore LW, Hancock AR, Chang SF, et al. Hepatitis B virus X protein and p53 tumor suppressor interactions in the modulation of apoptosis. *Proc Natl Acad Sci U S A*. 1997;94:14707–14712.
50. Wang HC, Huang W, Lai MD, Su IJ. Hepatitis B virus pre-S mutants, endoplasmic reticulum stress and hepatocarcinogenesis. *Cancer Sci*. 2006;97:683–688.
51. Caselmann WH. Transactivation of cellular gene expression by hepatitis B viral proteins: a possible molecular mechanism of hepatocarcinogenesis. *J Hepatol*. 1995;22(suppl):34–37.
52. Chang LJ, Pryciak P, Ganem D, Varmus HE. Biosynthesis of the reverse transcriptase of hepatitis B viruses involves de novo translational initiation not ribosomal frameshifting. *Nature*. 1989;337:364–368.
53. Bartenschlager R, Schaller H. Hepadnaviral assembly is initiated by polymerase binding to the encapsidation signal in the viral RNA genome. *EMBO J*. 1992;11:3413–3420.
54. Junker-Niepmann M, Bartenschlager R, Schaller H. A short cis-acting sequence is required for hepatitis B virus pregenome encapsidation and sufficient for packaging of foreign RNA. *EMBO J*. 1990;9:3389–3396.
55. Bartenschlager R, Junker-Niepmann M, Schaller H. The P gene product of hepatitis B virus is required as a structural component for genomic RNA encapsidation. *J Virol*. 1990;64:5324–5332.
56. Hirsch RC, Lavine JE, Chang LJ, Varmus HE, Ganem D. Polymerase gene products of hepatitis B viruses are required for genomic RNA packaging as well as for reverse transcription. *Nature*. 1990;344:552–555.
57. Pollack JR, Ganem D. An RNA stem-loop structure directs hepatitis B virus genomic RNA encapsidation. *J Virol*. 1993;67:3254–3263.
58. Rieger A, Nassal M. Specific hepatitis B virus minus-strand DNA synthesis requires only the 5' encapsidation signal and the 3'-proximal direct repeat DR1. *J Virol*. 1996;70:585–589.
59. Wang GH, Seeger C. The reverse transcriptase of hepatitis B virus acts as a protein primer for viral DNA synthesis. *Cell*. 1992;71:663–670.
60. Weber M, Bronsema V, Bartos H, Bosserhoff A, Bartenschlager R, Schaller H. Hepadnavirus P protein utilizes a tyrosine residue in the TP domain to prime reverse transcription. *J Virol*. 1994;68:2994–2999.
61. Zoulum F, Seeger C. Reverse transcription in hepatitis B viruses is primed by a tyrosine residue of the polymerase. *J Virol*. 1994;68:6–13.
62. Nassal M, Rieger A. A bulged region of the hepatitis B virus RNA encapsidation signal contains the replication origin for discontinuous first-strand DNA synthesis. *J Virol*. 1996;70:2764–2773.
63. Hu J, Boyer M. Hepatitis B virus reverse transcriptase and ϵ RNA sequences required for specific interaction in vitro. *J Virol*. 2006;80:2141–2150.
64. Gerlich WH, Robinson WS. Hepatitis B virus contains protein covalently attached to the 5' terminus of its complete DNA strand. *Cell*. 1980;21:801–809.
65. Lien JM, Aldrich CE, Mason WS. Evidence that a capped oligoribonucleotide is the primer for duck hepatitis B virus plus-strand DNA synthesis. *J Virol*. 1986;57:229–236.
66. Loeb DL, Hirsch RC, Ganem D. Sequence-independent RNA cleavages generate the primers for plus strand DNA synthesis in hepatitis B viruses: implications for other reverse transcribing elements. *EMBO J*. 1991;10:3533–3540.
67. Will H, Reiser W, Weimer T, et al. Replication strategy of human hepatitis B virus. *J Virol*. 1987;61:904–911.
68. Havert MB, Loeb DD. cis-acting sequences in addition to donor and acceptor sites are required for template switching during synthesis of plus-strand DNA for duck hepatitis B virus. *J Virol*. 1997;71:5336–5344.
69. Lutwick LI, Robinson WS. DNA synthesized in the hepatitis B Dane particle DNA polymerase reaction. *J Virol*. 1977;21:96–104.
70. Staprans S, Loeb DD, Ganem D. Mutations affecting hepadnavirus plus-strand DNA synthesis dissociate primer cleavage from translocation and reveal the origin of linear viral DNA. *J Virol*. 1991;65:1255–1262.
71. Yang W, Summers J. Illegitimate replication of linear hepadnavirus DNA through nonhomologous recombination. *J Virol*. 1995;69:4029–4036.
72. Yang W, Summers J. Integration of hepadnavirus DNA in infected liver: evidence for a linear precursor. *J Virol*. 1999;73:9710–9717.
73. Tuttleman JS, Pourcel C, Summers J. Formation of the pool of covalently closed circular viral DNA in hepadnavirus-infected cells. *Cell*. 1986;47:451–460.
74. Summers J, Smith PM, Horwich AL. Hepadnavirus envelope proteins regulate covalently closed circular DNA amplification. *J Virol*. 1990;64:2819–2824.
75. Jilbert AR, Wu TT, England JM, et al. Rapid resolution of duck hepatitis B virus infections occurs after massive hepatocellular involvement. *J Virol*. 1992;66:1377–1388.
76. Kajino K, Jilbert AR, Saputelli J, Aldrich CE, Cullen J, Mason WS. A study of woodchuck hepatitis virus infections revealed very rapid recovery after a prolonged viremia and infection of virtually every hepatocyte. *J Virol*. 1994;68:5792–5803.

77. Huovila AP, Eder AM, Fuller SD. Hepatitis B surface antigen assembles in a post-ER, pre-Golgi compartment. *J Cell Biol.* 1992;118:1305–1320.
78. Bonhoeffer S, Sniegowski P. Virus evolution: the importance of being erroneous. *Nature.* 2002;420:367–369.
79. Okamoto H, Imai M, Kametani M, Nakamura T, Mayumi M. Genomic heterogeneity of hepatitis B virus in a 54-year-old woman who contracted the infection through materno-fetal transmission. *Jpn J Exp Med.* 1987;57:231–236.
80. Orito E, Mizokami M, Ina Y, et al. Host-independent evolution and a genetic classification of the hepadnavirus family based on nucleotide sequences. *Proc Natl Acad Sci U S A.* 1989;86:7059–7062.
81. Mizokami M, Orito E, Ohba K, Ikeo K, Lau JY, Gojobori T. Constrained evolution with respect to gene overlap of hepatitis B virus. *J Mol Evol.* 1997;44:S83–S90.
82. Sterneck M, Gunther S, Gerlach J, et al. Hepatitis B virus sequence changes evolving in liver transplant recipients with fulminant hepatitis. *J Hepatol.* 1997;26:754–764.
83. Blackberg J, Kidd-Ljunggren K. Occult hepatitis B virus after acute self-limited infection persisting for 30 years without sequence variation. *J Hepatol.* 2000;33:992–997.
84. Norder H, Courouce AM, Coursaget P, et al. Genetic diversity of hepatitis B virus strains derived worldwide: genotypes, subgenotypes, and HBsAg subtypes. *Intervirology.* 2004;47:289–309.
85. Okamoto H, Tsuda F, Sakugawa H, et al. Typing hepatitis B virus by homology in nucleotide sequence: comparison of surface antigen subtypes. *J Gen Virol.* 1988;69:2575–2583.
86. Kramvis A, Kew M, Francois G. Hepatitis B virus genotypes. *Vaccine.* 2005;23:2407–2421.
87. Schaefer S. Hepatitis B virus: significance of genotypes. *J Viral Hepat.* 2005;12:111–124.
88. Thedja MD, Muljono DH, Nurainy N, Sukowati CH, Verhoef J, Marzuki S. Ethnogeographical structure of hepatitis B virus genotype distribution in Indonesia and discovery of a new subgenotype, B9. *Arch Virol.* 2011;156:855–868.
89. Mulyanto, Depamede SN, Wahyono A, et al. Analysis of the full-length genomes of novel hepatitis B virus subgenotypes C11 and C12 in Papua, Indonesia. *J Med Virol.* 2011;83:54–64.
90. Meldal BH, Moula NM, Barnes IH, Boukef K, Allain JP. A novel hepatitis B virus subgenotype, D7, in Tunisian blood donors. *J Gen Virol.* 2009;90:1622–1628.
91. Devesa M, Loureiro CL, Rivas Y, et al. Subgenotype diversity of hepatitis B virus American genotype F in Amerindians from Venezuela and the general population of Colombia. *J Med Virol.* 2008;80:20–26.
92. Tran TT, Trinh TN, Abe K. New complex recombinant genotype of hepatitis B virus identified in Vietnam. *J Virol.* 2008;82:5657–5663.
93. Tatematsu K, Tanaka Y, Kurbanov F, et al. A genetic variant of hepatitis B virus divergent from known human and ape genotypes isolated from a Japanese patient and provisionally assigned to new genotype J. *J Virol.* 2009;83:10538–10547.
94. Simmonds P, Midgley S. Recombination in the genesis and evolution of hepatitis B virus genotypes. *J Virol.* 2005;79:15467–15476.
95. Tabor E, Purcell RH, London WT, Gerety RJ. Use of and interpretation of results using inocula of hepatitis B virus with known infectivity titers. *J Infect Dis.* 1983;147:531–534.
96. Chu CJ, Lok ASF. Clinical significance of hepatitis B virus genotypes. *Hepatology.* 2002;35:1274–1276.
97. Tanaka Y, Mizokami M. Genetic diversity of hepatitis B virus as an important factor associated with differences in clinical outcomes. *J Infect Dis.* 2007;195:1–4.
98. Guettouche T, Hnatyszyn HJ. Chronic hepatitis B and viral genotype: the clinical significance of determining HBV genotypes. *Antivir Ther.* 2005;10:593–604.
99. Lin CL, Kao JH. The clinical implications of hepatitis B virus genotype: recent advances. *J Gastroenterol Hepatol.* 2011;26(suppl 1):123–130.
100. Datta S. An overview of molecular epidemiology of hepatitis B virus (HBV) in India. *Virol J.* 2008;5:156.
101. Wang HP, Rogler CE. Topoisomerase-I mediated integration of hepadnavirus DNA in vitro. *J Virol.* 1991;65:2381–2392.
102. Su TS, Lai CJ, Huang JL, et al. Hepatitis B virus transcript produced by RNA splicing. *J Virol.* 1989;63:4011–4018.
103. Gunther S. Genetic variation in HBV infection: genotypes and mutants. *J Clin Virol.* 2006;36:S3–S11.
104. Kramvis A, Kew M. The core promoter of hepatitis B virus. *J Viral Hepat.* 1999;6:415–427.
105. Kuang SY, Jackson PE, Wang JB, et al. Specific mutations of hepatitis B virus in plasma predict liver cancer development. *Proc Natl Acad Sci U S A.* 2004;101:3575–3580.
106. Su IJ, Wang HC, Wu HC, Huang WY. Ground glass hepatocytes contain pre-S mutants and represent preneoplastic lesions in chronic hepatitis B virus infection. *J Gastroenterol Hepatol.* 2008;23:1169–1174.
107. Pawlotsky JM. The concept of hepatitis B virus mutant escape. *J Clin Virol.* 2005;34(suppl 1):S125–S129.
108. Torresi J. The virological and clinical significance of mutations in the overlapping envelope and polymerase genes of hepatitis B virus. *J Clin Virol.* 2002;25:97–106.
109. Carman WF, Zanetti AR, Karayiannis P, et al. Vaccine-induced escape mutant of hepatitis B virus. *Lancet.* 1990;336:325–329.
110. Cooreman MP, Leroux-Roels G, Paulij WP. Vaccine- and hepatitis B immune globulin-induced escape mutations of hepatitis B virus surface antigen. *J Biomed Sci.* 2001;8:237–247.
111. Hoofnagle JH, Doo E, Liang TJ, Fleischer R, Lok AS. Management of hepatitis B: summary of a clinical research workshop. *Hepatology.* 2007;45:1056–1075.
112. Stuyver LJ, Locarnini SA, Lok A, et al. Nomenclature for antiviral-resistant human hepatitis B virus mutations in the polymerase region. *Hepatology.* 2001;33:751–757.
113. Angus P, Vaughan R, Xiong S, et al. Resistance to adefovir dipivoxil therapy associated with development of a novel mutation in the HBV polymerase. *Gastroenterology.* 2003;125:292–297.
114. Locarnini S, Yuen L. Molecular genesis of drug resistant and vaccine escape HBV mutants. *Antivir Ther.* 2010;15:451–461.
115. Kim KH, Kim ND, Seong BL. Discovery and development of anti-HBV agents and their resistance. *Molecules.* 2010;15:5878–5908.
116. Datta S, Banerjee A, Chandra PK, et al. Analysis of hepatitis B virus X gene phylogeny, genetic variability and its impact on pathogenesis: implications in Eastern Indian HBV carriers. *Virology.* 2008;382:190–198.