# Origins and Evolution of Hepatitis B Virus and Hepatitis D Virus

**Margaret Littlejohn, Stephen Locarnini, and Lilly Yuen**

Molecular Research and Development, Victorian Infectious Diseases Reference Laboratory, Doherty Institute, Melbourne 3000, Australia

*Correspondence:* stephen.locarnini@mh.org.au

Cold Spring Harbor Perspectives in Medicine
www.perspectivesinmedicine.org

Members of the family Hepadnaviridae fall into two subgroups: mammalian and avian. The detection of endogenous avian hepadnavirus DNA integrated into the genomes of zebra finches has revealed a deep evolutionary origin of hepadnaviruses that was not previously recognized, dating back at least 40 million and possibly >80 million years ago. The non-primate mammalian members of the Hepadnaviridae include the woodchuck hepatitis virus (WHV), the ground squirrel hepatitis virus, and arctic squirrel hepatitis virus, as well as a number of members of the recently described bat hepatitis virus. The identification of hepatitis B viruses (HBVs) in higher primates, such as chimpanzee, gorilla, orangutan, and gibbons that cluster with the human HBV, as well as a number of recombinant forms between humans and primates, further implies a more complex origin of this virus. We discuss the current theories of the origin and evolution of HBV and propose a model that includes cross-species transmissions and subsequent recombination events on a genetic backbone of genotype C HBV infection. The hepatitis delta virus (HDV) is a defective RNA virus requiring the presence of the HBV for the completion of its life cycle. The origins of this virus remain unknown, although some recent studies have suggested an ancient African radiation. The age of the association between HDV and HBV is also unknown.

The long-term evolutionary history of the hepatitis B virus (HBV) remains elusive and controversial. Numerous theories have been proposed for the origin of HBV (Simmonds 2001), but they do not adequately explain the current geographical distribution of the HBV genotypes (Fig. 1). These theories are predominantly derived from molecular phylogenetic analyses of the HBV genome, but the timescales inferred by these methods are strongly influenced by the parameters and statistical algorithms used during the inferencing process, with a key parameter being the mutation rate of the organism. Because the HBV reverse transcriptase (RT) lacks proofreading ability and allows nucleotide misincorporation during genome replication, HBV is expected to have a high mutation rate. Indeed, commonly accepted rates for HBV are $\sim 2.0 \times 10^{-5}$ nucleotide substitutions per site per year (Orito et al. 1989). However, this rate inferred that the most recent common ancestor (MRCA) for human HBVs is ~2000 to 3000 years old. It could be possible that HBV genotypes have emerged within this time frame, but it is insufficient to allow for the complete global expansion of HBV infections in >350 million chronically infected humans and primates.

Editors: Christoph Seeger and Stephen Locarnini
Additional Perspectives on Hepatitis B and Delta Viruses available at www.perspectivesinmedicine.org

 doi: 10.1101/cshperspect.a021360
Cite this article as *Cold Spring Harb Perspect Med* 2016;6:a021360

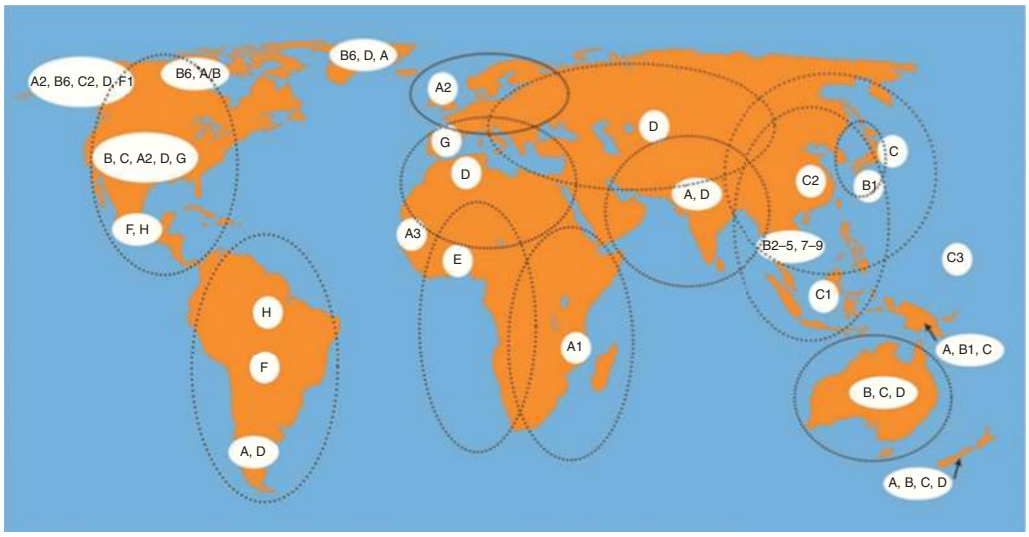


**Figure 1.** Geographical distribution of the HBV genotypes and subgenotypes. Genotype I and J are not shown as they have not been ratified by the International Committee on Taxonomy of Viruses (ICTV); genotype I is found in Southern China and Vietnam, whereas genotype J was identified from a Japanese soldier who had spent time in Borneo. Subgenotype C3 in the Western Pacific is referring to Melanesian areas, such as the Solomon Islands and Vanuatu. (From Spradling et al. 2013; with permission from the authors.)

The disparity between the inferred MRCA of human HBV and the current geographical distribution may also be explained by the unique features of the HBV itself. To elucidate the evolutionary history of HBV, it is necessary to consider the complex genome structure of this virus. The HBV genome is mainly composed of overlapping open reading frames (ORFs), which limit nonsynonymous mutations, and a significant proportion of the single coding regions are either embedded within regulatory elements or are involved in secondary structure interactions. This compact organization of the viral genome places considerable constraints on where nucleotide substitutions can occur, and suggests a substantially slower long-term mutation rate.

In contrast, far fewer theories have been proposed for the origin and evolution of the hepatitis D virus (HDV). Given that the HDV genome is composed of a viroid-like component and a protein-coding component, a plausible theory is that the viral genome was the product of a recombination event between RNA molecules derived from different sources (Taylor 2014). However, recent revelations concerning host cellular RNA, including the recognition of circular RNA, have led to new models and hypotheses on this topic.

In this review, theories on the possible origins of human HBV and HDV will be discussed, as well as the evolutionary processes involved with emergence of the current 10 known major genotypes of HBV and eight genotypes of HDV.

## THE FAMILY HEPADNAVIRUSES

The human HBV is the prototype member of the family Hepadnaviridae. Detailed information about the virus can be found in Burns and Thompson (2014). The Hepadnaviridae family comprises two genera: *Orthohepadnaviruses* that infect mammals, and *Avihepadnaviruses* that infect avian species (Table 1). The phylogenetic relatedness between the Hepadnaviridae species is shown in Figure 2.

## VIRAL GENOMIC FOSSILS

The ability for hepadnaviruses to undergo endogenization into the host genome has only

**Table 1.** Comparison of the different members of the Hepadnaviridae family

| | *Orthohepadnaviridae* | | | *Avihepadnaviridae* |
|---|---|---|---|---|
| | HBV | WHV | BtHV | DHBV |
| **Physical characteristics** | | | | |
| Genome size | 3182–3284 | 3308–3320 | 3149–3377 | 3021–3027 |
| nt Homology with HBV | 100 | 70 | 48 | 40 |
| Number of ORFs | 4 | 4 | 4 | 4 |
| **ORF—amino acid lengths** | | | | |
| Pre-S | 163 | 204–205 | 159–224 | 161–163 |
| S | 226 | 222 | 223–224 | 167 |
| Pre-C | 29 | 30 | 28–33 | 43 |
| C | 183 | 187 | 188–189 | 262 |
| P | 838 | 879 | 827–902 | 786–788 |
| X | 145 | 141 | 135–144 | 114 |

HBV, hepatitis B virus; WHV, woodchuck hepatitis virus; BtHV, bat hepatitis virus; DHBV, duck hepatitis B virus; nt, nucleos(t)ides; ORFs, open reading frames; Pre-S, presurface; S, surface; pre-C, precore; C, core; P, polymerase.

been realized in recent years. The first evidence of HBV-derived endogenous viral elements (EVEs) was identified by BLAST-based analyses on the zebra finch genome, although this bird species is not productively infected with a hepatitis virus today (Gilbert and Feschotte 2010). Gilbert and Feschotte (2010) identified at least 12 HBV-derived endogenous sequences, extending over 4%–40% of the "modern-day" duck HBV (DHBV) genome, and labeled them endogenous zebra finch HBVs (eZHBVs). Phylogenetic analysis suggested the integration events occurred at least 19 million years ago (mya). The study also identified orthologous insertions in other bird genomes of the order Passeriformes, including other finch species (Estrildidae family), the olive sunbird (Nectariniidae family), and the dark-eyed junco (Emberizidae family) (Gilbert and Feschotte 2010). Subsequent to this landmark study, similar avihepadnaviral elements have been identified in the genome of budgerigars of the order Psittaciformes (Cui and Holmes 2012). Comparative phylogenetic analysis on these endogenous and existing exogenous avian hepadnaviruses established the eZHBVs are far more divergent than modern-day DHBVs and form their own distinct lineages (Gilbert and Feschotte 2010), indicative of multiple or recurrent genomic integration events. A more recent study reported a genomic record of Hepadnaviridae endogenizations estimated to have occurred during a period of bird evolution from 12 to 82 mya (Suh et al. 2013), and constituted the first discovery of a Mesozoic paleovirus genome (>82 mya). Not surprisingly then, birds have been proposed as the ancestral hosts of Hepadnaviridae, and mammalian HBVs probably emerged after a bird–mammal host switch (Suh et al. 2013). These paleovirological studies support a reevaluation of current theories on hepadnaviral origins and subsequent evolution.

## HBV AS A ZOONOTIC INFECTION

### Bat Hepadnaviruses

The recent identification of the bat hepatitis virus (BtHV) has substantially changed the perception of the zoonotic potential of HBV (Drexler et al. 2013; He et al. 2013). The proportion of BtHV DNA-positive bats ranged from 2.2% to 9.3% in the various bat species (Drexler et al. 2013; He et al. 2013). There are >1100 species of bats (*Chiroptera*) in the world, representing ~25% of all known species of mammals on Earth (Drexler et al. 2011). Multiple factors, such as longevity, migratory activity, large and dense roosting communities, and close social interaction, predispose bats as reservoirs for a wide diversity of viruses (Calisher et al. 2006). This "viral burden" or "virome" in bats can have considerable public health signifi-

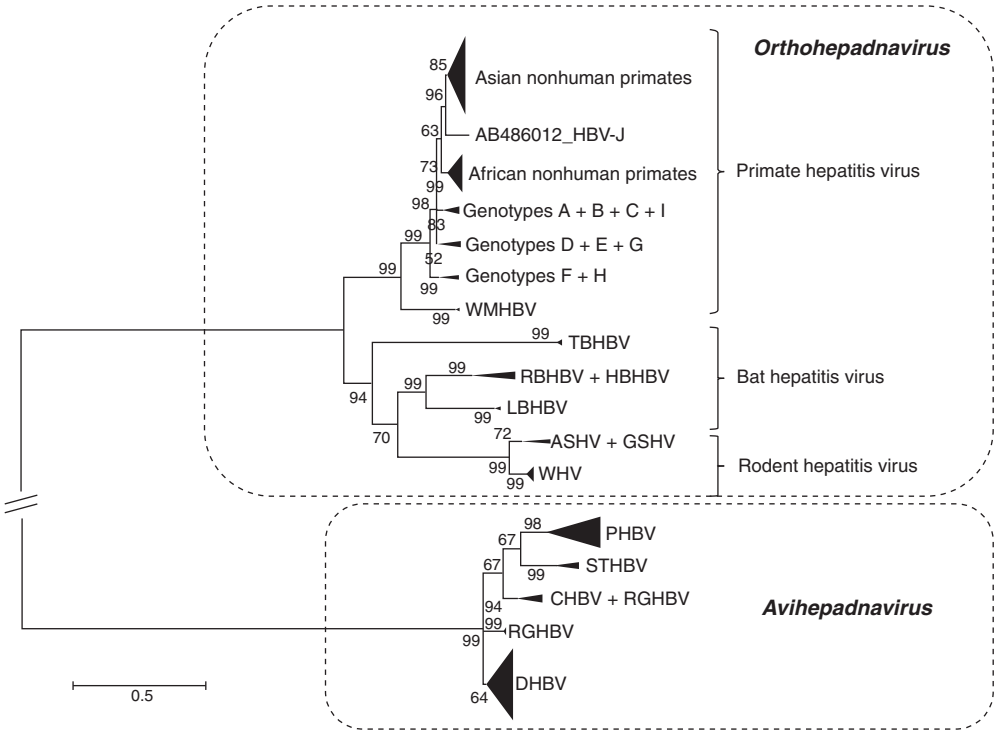


**Figure 2.** Maximum likelihood (ML) phylogenetic tree showing the genetic relationships between all members of the Hepadnaviridae family. These include 219 complete hepatitis B virus (HBV) genomes determined from primate, rodent, bat, and avian species. Primate HBV sequences were from humans (genotypes A–J), nonhuman primates from Asia (gibbons and orangutans), apes from Africa (chimpanzees and gorilla), and woolly monkey hepatitis B virus (WMHBV). Rodent HBV sequences were from ground squirrel hepatitis virus (GSHV), arctic squirrel hepatitis virus (ASHV), and woodchuck hepatitis virus (WHV), whereas the bat HBV sequences were from Burmese long-fingered bats (LBHBV), African roundleaf bats (RBHBV), African horseshoe bats (HBHBV), and tent-making bats (TBHBV) from Panama. The avihepadnavirus HBV sequences were determined from duck hepatitis B virus (DHBV), Ross goose hepatitis virus (RGHBV), crane hepatitis B virus (CHBV), heron hepatitis B virus (HHBV), stork hepatitis B virus (STHBV), and parrot hepatitis B virus (PHBV). The ML tree was bootstrapped 1000× and only values >50% are shown.



cance (Drexler et al. 2012). Although the transmission mode for the majority of bat-borne viruses has not been elucidated, they can readily be passed to other mammals resulting in outbreaks of epidemic and life-threatening proportions (Dobson 2005; Chu et al. 2008). Ongoing studies are in progress to determine the tropism and prevalence of BtHVs in other bat species, as well as their possible zoonotic potential. It should be noted that the PreS1 protein of BtHV contains the motif [NPLGF(F/L)] (Ni et al. 2014), which enables human HBV to bind to the hepatocyte receptor sodium taurocholate cotransporting protein (NTCP) and gain entry into liver cells (Yan et al. 2012). In vitro studies have also confirmed that BtHV has the potential to infect human hepatocytes, thereby posing a very real public health threat, particularly in bat-infested areas (Drexler et al. 2013).

The genomic structure of BtHV is similar to the other orthohepadnaviruses (Table 1), and is similar in size to the primate HBVs and rodent hepatitis viruses. Although the genomic diversity of BtHV genomes to all known hepadnaviruses is >35%, it should be noted that the BtHVs isolated from the four bat species to date also have genomic diversities of up to 39% between each other (Fig. 2) (Drexler et al. 2013),

Cold Spring Harbor Perspectives in Medicine
www.perspectivesinmedicine.org

suggesting that BtHV originated in bats a long time ago, possibly during the early phases of bat evolution, and the different species of BtHVs have since coevolved with their host species.

### Rodent and Avian Hepadnaviruses

Although rodent and avian HBVs cannot infect primary human hepatocytes in vitro, it is known that infectious cDNA constructs from these HBVs can replicate, assemble, and release infectious virus when transfected into human hepatocytes (Schlicht et al. 1987; Seeger et al. 1989). Thus, these groups of viruses are unlikely to pose a zoonotic threat to man. Nonetheless, the proposal discussed above that members of *Avihepadnaviridae* made a zoonotic jump to mammals (Suh et al. 2013) will require further experimentation to validate in terms of specific virus-cell receptors, such as NTCP (Yan et al. 2012).

### Nonhuman Primate Hepadnaviruses

Natural infections with species-specific Hepadnaviridae have been well documented for gorillas, chimpanzees, orangutans, gibbons, and woolly monkeys (Norder et al. 1996; Lanford et al. 1998; Warren et al. 1999; Grethe et al. 2000; Robertson and Margolis 2002; Sall et al. 2005), and there is a strong association between their genetic relatedness and the natural habitat locations of these primates (Starkman et al. 2003). For example, gorilla HBV genomes are most closely related to those of chimpanzee HBVs from *Pan troglodytes troglodytes*, and their habitats are located in overlapping geographical ranges (Starkman et al. 2003), whereas HBV isolated from chimpanzee species in other regions are genetically more distantly related. This relationship is comparable to the current geographical distribution of human HBV genotypes. More interestingly, phylogenetic analysis performed on the genomes of all primate HBVs has revealed that both human and nonhuman primates do belong to the same lineage, and no distinct MRCAs could be inferred for the two primate groups (see Fig. 2).

The nonhuman primate HBVs share the consensus PreS1 sequence that allows binding to the NTCP receptor on human hepatocytes (Ni et al. 2014). Thus, drawing from the evolutionary history of HIV-1 and HIV-2 that were initiated from cross-species transmission of simian immunodeficiency virus (SIV) from infected chimpanzees, gorillas, and macaques into humans (Sharp and Hahn 2011), a similar scenario may have occurred between the nonhuman and human primate hepadnaviruses. Determination of the timing of these events could provide important insights into HBV origins and subsequent evolution.

## HUMAN HBV

### Role of HBV Genotype and Subgenotypes

Classification of human HBVs is based on phylogenetic analysis of the complete viral genome (Miyakawa and Mizokami 2003; Norder et al. 2004; Olinger et al. 2008; Tatematsu et al. 2009). The convention is to group HBVs with up to 8% nucleotide divergence in their complete genome sequences into genotypes, and intragenotypic HBVs with 4%–8% nucleotide divergence into subgenotypes (Okamoto et al. 1988; Norder et al. 1994; Sugauchi et al. 2001). Human HBVs are currently classified into 10 genotypes (A-J) and a growing number of subgenotypes.

It was evident soon after the classification scheme was introduced in 1988 that HBV genotypes have distinct geographical distributions (Okamoto et al. 1988), and this association formed the base from which the majority of HBV evolutionary theories were constructed (see section on Theories of HBV Origins). The current global distribution of HBV genotypes is shown in Figure 1, and more detail on the topic can be found in a review by Kramvis et al. (2005a) (genotypes A–H) and in a number of other study reports (genotypes I and J) (Seeger et al. 1989; Norder et al. 2004; Tatematsu et al. 2009). Interestingly, evidence is emerging that HBV subgenotypes also have distinct geographical distributions as shown in Figure 1, implying an important role in the evolutionary history of HBV. For example, phylogenetic analysis on the complete HBV genome sequences showed that genotypes A–D and F can be further divided into subgenotypes (identified by numbers). Some subgenotypes, such as D1, D2, and D3, are

widely distributed (Kimbi et al. 2004; Banerjee et al. 2006), whereas others show distinct geographical localization. For instance, A1 is predominantly found in Africa and Asia, whereas A2 is mainly found in Europe and North America (Sugauchi et al. 2004a); B1 is confined to Japan, whereas B2 (a genotype B and C recombinant) is found in the rest of Asia (Sugauchi et al. 2004b); and C1 is the dominant strain in South and Southeast Asia, whereas C2 is mainly found in North Asia (Huy et al. 2004; Chan et al. 2005). By using a variety of phylogenetic methods, Kramvis and Paraskevis (2013) were able to show HBV-A1 originated in Africa, and that its presence in Asia and Latin America was the consequence of the slave trade operating from the 9th to 19th centuries. Conversely, the absence of subgenotypes in HBV genotype E has been assumed to be the consequence of its more recent genesis (Mulders et al. 2004; Fujiwara et al. 2005; Kramvis et al. 2005b; Huy et al. 2006). This is supported by the observation that HBV-E has not been isolated from Americans of African origin, who are mainly descendants of African slaves from Venezuela and Brazil (Quintero et al. 2002; Motta-Castro et al. 2005).

## Mutation Rate of the Hepadnavirus Genome

An essential factor for calculating the evolutionary divergence of hepadnaviruses is the inferred mutation rate of the virus genome. It is defined as the number of base substitutions within the genome per site per year of continuous virus replication in the host. HBV mutation rates are often inferred by comparing the HBV sequences either from mothers and children of maternally acquired cases, or chronic carriers over a specific time frame. However, these mutation rates are considered short term, and should be used with caution. There are two important observations to support this claim. First, mutation rates can vary by at least one order of magnitude if the HBV genome sequences used for inferencing were from individuals in the HBeAg-negative phase (anti-HBe-positive) of disease instead of being in the asymptomatic HBeAg-positive phase ($10^{-4}$ vs. $10^{-5}$ substitutions per site per year [s/s/y], respectively) (Harrison et al. 2011). Second, HBV mutations have the tendency to "revert" back to the genotype consensus over time ("genotypic self-reversion"), thus, suggesting that the HBV genome would not have changed significantly over time despite having a high mutation rate ("The Red Queen hypothesis") (Tedder et al. 2013). These observations further support the model that HBV has a long evolutionary history and may have coevolved with humans in the ancient past.

The fortuitous finding of an HBV-infected mummified child from 16th century Korea provided a unique opportunity to recalculate the mutation rate of HBV (Kahila Bar-Gal et al. 2012). Phylogenetic analysis of this ancient HBV (aHBV) confirmed the child was infected with HBV-C2 (Ahn et al. 2010). Using a range of mutation rates not slower than $10^{-6}$ s/s/y (Orito et al. 1989, 2001; Zhou and Holmes 2007), Korean HBV was estimated to have originated sometime between 3000 and 100,000 years ago. This timeframe conclusively places HBV in East Asia at least 3000 years ago (Orito et al. 1989; Jazayeri et al. 2010). However, nucleotide divergence analysis surprisingly revealed that the aHBV genome had up to 97% identity with other modern HBV-C2 genome sequences from Korea, despite being isolated from a child who had lived ~350 years previously. This finding strongly supports the hypothesis that the long-term mutation rate of HBV is substantially slower ($10^{-6}$–$10^{-7}$ s/s/y) than the currently accepted range ($10^{-4}$–$10^{-5}$ s/s/y).

## Mixed Infections and the Role of Recombination

Double infection with two different HBV genotypes was first shown using serological methods (Hess et al. 1977; Tabor et al. 1977). The advances made in genotyping methodologies have enabled double infections of HBV, with different genotypes, to be detected in patients with chronic hepatitis B (Kao et al. 2001). Numerous studies have since described high rates of such double infections with different HBV genotypes, ranging from 4.0% to 18.0% (Ding et al. 2003; Kato et al. 2003; Osiowy and Giles 2003; Chen et al. 2004; Olinger et al. 2006).

Coinfection or even superinfection with two different HBV genotypes in the same host may enable exchange and recombination of genetic materials between two viral strains to take place. There are many opportunities throughout the HBV life cycle for this to occur, including RNA–RNA during packaging of pregenomic RNA, as well as DNA–DNA during reverse transcription and generation of double-stranded linear genomes (Yang and Summers 1995). Opportunities are also present during RNA transcription from the two main subpopulations of minichromosomes for recombination events to occur (Newbold et al. 1995).

Some HBV recombinants are now endemic in certain geographic regions, and have been assigned their own genotype or subgenotype. Most notable is the recombination between isolates of HBV genotypes B and C (Sugauchi et al. 2002), resulting in HBV subgenotype B2, which contains the precore and core genes of an HBV genotype C, and is found throughout Asia (Sugauchi et al. 2002). The parent subgenotype B1 is not a recombinant and is mainly confined to Japan. Subgenotype HBV genotype I can be regarded as a triple recombinant containing elements from genotypes A, G, and C, and is predominantly found in Vietnam (Huy et al. 2008), whereas genotype J appears to represent a recombinant between genotype C and gibbon HBV (Fig. 3A) (Tatematsu et al. 2009). Other recognized recombinants that localize to specific regions have also been described, including the C/D-recombinant circulating in Tibet (Cui et al. 2002; Wang et al. 2005; Zeng et al. 2005).

The HBV genome has a very complex structure with the majority of genes located in overlapping reading frames, and viral regulatory elements embedded within coding regions. Thus, HBV genomes are expected to have alternating regions of variable and conserved domains. Bowyer and Sim (2000) referred to these domains as mosaic structures, and showed that some of the more variable regions were associated with recombination events rather than accumulation of point mutations. It is possible that modern HBV genomes comprise allelic modules with different properties caused by evolutionary selection pressure on key regions, driven by host–virus relationship and pathogenesis of chronic hepatitis B. Simmonds and Midgley (2005) have also described the HBV genome as modular, representing a blend of small segments from genomes of different human and primate HBV strains. Given the opportunity, key antigenic epitopes in the envelope and core proteins that are associated with immune escape, as well as binding sites for transcription factors that control viral replication (Fischer et al. 2006), can rapidly emerge via recombination events within the genome of HBVs in endemic countries.

Simmonds and Midgley (2005) and Yang and colleagues (2006) have tested for the presence of recombination in the HBV genome using tree-order scanning. Most (90%) of the recombinants detected were either B/C or A/D hybrids. Other recombinants identified included A/B/C, A/C, A/E, A/G, C/D, C/F, C/G, C/U (U = unknown genotype) and B/C/U hybrids. Genotypes A and C tend to show a higher tendency for recombination than other genotypes, and up to eight breakpoint hotspots have now been mapped as crossover points for intergenotype recombination (Fig. 3B). Interspecies recombination events between human and chimpanzee HBV (Magiorkinis et al. 2005), as well as the human and gibbon HBV (Fig. 3A) (Tatematsu et al. 2009) discussed previously, have also been described. The former study characterized an HBV-recombinant that was isolated from a wild-caught chimpanzee (*Pan troglodytes schweinfurthii*) in East Africa (Vartanian et al. 2002). The genome of this isolate (FG) contained a segment from human HBV subgenotype C2, supporting the occurrence of recombination among various primate HBV strains, which may have played an important role in the current distribution of HBV genotypes and subgenotypes following *Homo sapiens* exit from Africa and subsequent global diaspora (see section on HBV and Indigenous Populations).

## THEORIES OF HBV ORIGINS

At least five theories have been proposed on the origins of HBV.


**Figure 3.** Diagrammatic representations of recombination among hepatitis B viruses (HBVs). (*A*) Gibbon HBV and a human genotype C HBV, recombining in the core gene to generate the genotype J. (*B*) Modular (at least eight) nature of the HBV genome recognized from common HBV recombinants identified to date (Simmonds and Midgley 2005; Bowyer and Sim 2000). Simmonds and Midgley (2005) have identified each one of these as a major phylogenetic discontinuity, which they propose has led to the emergence of the major HBV genotypes recognized today.

1. New World origin (out of South America): This account suggests that human HBV had originated in the Americas and was brought into the Old World during the last 400 years after contact with Europeans (Bollyky et al. 1997). This theory is difficult to reconcile given the current geographical spread of global HBV genotypes, and the presence of HBV in nonhuman primate species. The isolation of the aHBV from the Korean mummy dating from the 16th century (Kahila Bar-Gal et al. 2012) also strongly contradicts this theory.

2. Cospeciation: A second theory suggests that Hepadnaviridae have evolved in parallel to specific species of primates over the past 10–35 million years. This theory is supported by the times of divergence inferred between the different ortho- (mammals) and avi- (birds) Hepadnaviridae being similar to those inferred between the respective hosts based on phylogeny as well as fossil-based evidence (MacDonald et al. 2000). Additionally, chimpanzees of different geographical regions have chimpanzee hepatitis B virus (ChHBV) with genome sequences that constitute unique phylogenetic clades, thus, further supporting HBV cospeciation, at least in chimpanzees (Hu et al. 2001). However, the high level of similarity between the HBV genome sequences from gorillas and chimpanzees, in particular from *P. t. troglodytes* with whom gorillas share an overlapping geographical range, poses a problem for this interpretation.

3. Coevolution as anatomically modern humans (AMH) migrated out of Africa: An alternative theory is that human HBV has coevolved with AMH since their migration out of Africa ~100,000 years ago (Norder et al. 1994; Magnius and Norder 1995). Although this theory can explain the current geographical spread of HBV genotypes, it does not fit with the close genetic relationships observed between primate and human HBV. Another inconsistency is that Native Americans predominantly have genotype F infections, whereas northeast-Asians, who are their closest relatives genetically, have genotypes B and C infections. Nonetheless, support for this "out of Africa" theory is found in a recent study by Paraskevis and colleagues using Bayesian phylogenetic analysis of hepatitis B surface antigen gene sequences, who concluded HBV "jumped" into humans between 22,000 and 47,100 years ago from an unknown source, and proposed that humans then went on to infect nonhuman primates in Africa, Asia, and the New World (Paraskevis et al. 2013).

4. Cross-species transmission: A theory that explains the close relationship between human and nonhuman primate HBVs proposes several instances of cross-species transmission. This is supported by the finding that geographical regions with potentially high HBV transmission rates between human and nonhuman primates are also areas of high HBV prevalence among humans (i.e., Southeast Asia and Africa). There are several reports of known cross-species transmission cases, such as human HBV identified from chimpanzees (Hu et al. 2000; Takahashi et al. 2000), the identification of human/primate HBV recombinants (Magiorkinis et al. 2005; Simmonds and Midgley 2005), infection of Mauritian macaques with human HBV (Dupinay et al. 2013), and a gibbon variant isolated from a chimpanzee (Grethe et al. 2000), which would lend substantial credence to this model.

5. Bat origin: A theory put forward recently following the detection of Hepadnaviridae in several bat species (Drexler et al. 2013) suggests a bat origin of primate hepadnaviruses. This theory could explain the unexpected absence of HBV detection in other primate species, including cercopithecoid monkeys and other non-Simiiformes monkeys.

Given the arguments for and against each of these five theories, it is probable that HBV evolution cannot be explained by any single theory. The reality probably involves cospecies evolution within birds, rodents, and bats, followed by a series of cross-species transmission events to explain the close relationship between human and nonhuman primate HBVs observed today. Challenges for any unifying theory include the high level of genome divergence observed between HBV sequences of New World woolly monkeys and other nonhuman primates, which cannot be explained by the cross-species transmission theory, and also that HBV has only been detected in rodent species of the New World. If HBV coevolved with avian, rodent, and primate species, then why is it not found in all rodent and primate species? In addition, if HBV emerged out of Africa with AMH, then why are people from the New World, who are

genetically most closely related to humans in the Far East, predominantly infected with HBV genotypes F and H rather than the genetically unrelated HBV genotypes B and C that are found in the Far East?

## THE HUMAN FAMILY TREE: ADMIXTURES AND ARCHAIC HUMAN DNA

All evidence discussed thus far supports a long and complex evolutionary history for HBV, and its origin could even date back to the Cretaceous period when birds first began to emerge (see section on Viral Genomic Fossils). An important missing element in all of the theories proposed on primate HBV origin to date is the influence of archaic hominids on the evolution of this virus.

The sequencing of two archaic human genomes (Neanderthal [Green et al. 2010] and Denisovan [Reich et al. 2010]) has resulted in a reappraisal of human evolution. Examination of the two ancient DNA sets revealed some surprising admixtures, with up to 4% of Neanderthal DNA found in the nuclear DNA of modern Europeans and Asians (with none in Africans) and up to 5% of Denisovan DNA in Melanesians from Papua New Guinea and the Bougainville Islands. Importantly, no Denisovan DNA has yet been detected in the genome of Neanderthals or other living humans. Additionally, using the DNA sequences from 61 loci of modern African genomes, Hammer and colleagues (2011) inferred that ~2% of *H. sapiens* DNA originated ~35 thousand years ago (kya) from an archaic population that had split from the ancestors of AMH about 700 kya. Taken together, these results suggest interbreeding between ancestors of *H. sapiens* and archaic humans on at least three occasions (Fig. 4). This AMH origin model has been termed "leaky replacement," as opposed to the complete sequential replacement of the archaic population model (Gibbons 2011). The influence of these various groups of archaic humans on the evolutionary history of HBV would be difficult

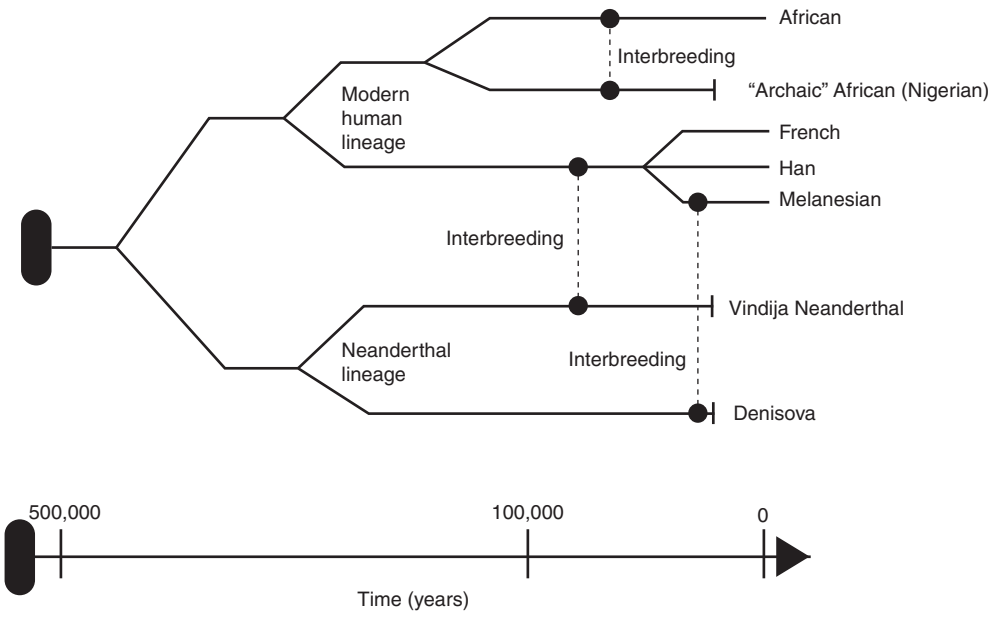


**Figure 4.** The human family tree updated with the recent recognition of admixtures between the common *Homo sapiens* lineage with an archaic African (Nigerian) as well as with the Neanderthal and Denisovan lineage. This has become known as the "leaky replacement model" of anatomically modern human (AMH) origins. (From Gibbons 2011; modified, with permission, from the author.)

to decipher. However, the possibility that human HBV may have originated, at least in part, from these archaic humans should not be discounted.

## HBV AND INDIGENOUS POPULATIONS: IMPLICATIONS FOR HBV ORIGINS

Present-day indigenous Australians are likely to be one of the oldest continuous populations of AMH outside Africa (Rasmussen et al. 2011). Comparative analysis between the genomes of an indigenous Australian who lived 100 years ago in southern Western Australia and 1220 living individuals of 79 populations, confirmed that indigenous Australians are descendants of the early wave (presumably first) of human dispersal into Eastern Asia (Sunda) possibly 62,000 to 75,000 years ago (Fig. 5) (Rasmussen et al. 2011). This study further showed that this dispersal was separate from the wave that gave rise to modern Europeans and Asians 25,000 to 38,000 years ago. This two-wave model of *H. sapiens* migration proposes that the ancestors of indigenous Australians and related populations (Melanesians) had diversified from the Eurasian population before their split into European and Asian populations (Rasmussen et al. 2011). If these archaic populations were infected with HBV at a similar endemicity as seen in the indigenous populations of today, then it may be possible to hypothesize an alternative origin of HBV genotypes, reflecting their geographical distribution. Molecular phylogenetic studies of HBV genotypes repeatedly show that the genotype C viruses appear to be the oldest of the human HBVs (Paraskevis et al. 2013) and their HBsAg is the closest of all the other human HBVs to the primate hepatitis viruses (Norder et al. 2004). It is important to note then that indigenous Australians are infected with a unique HBV subgenotype C4, which may very well represent the oldest strain of HBV infecting AMH (Davies et al. 2013; Littlejohn et al. 2014).

Although the geographical distribution of HBV genotypes and subgenotypes does provide some insights into HBV evolutionary relation-

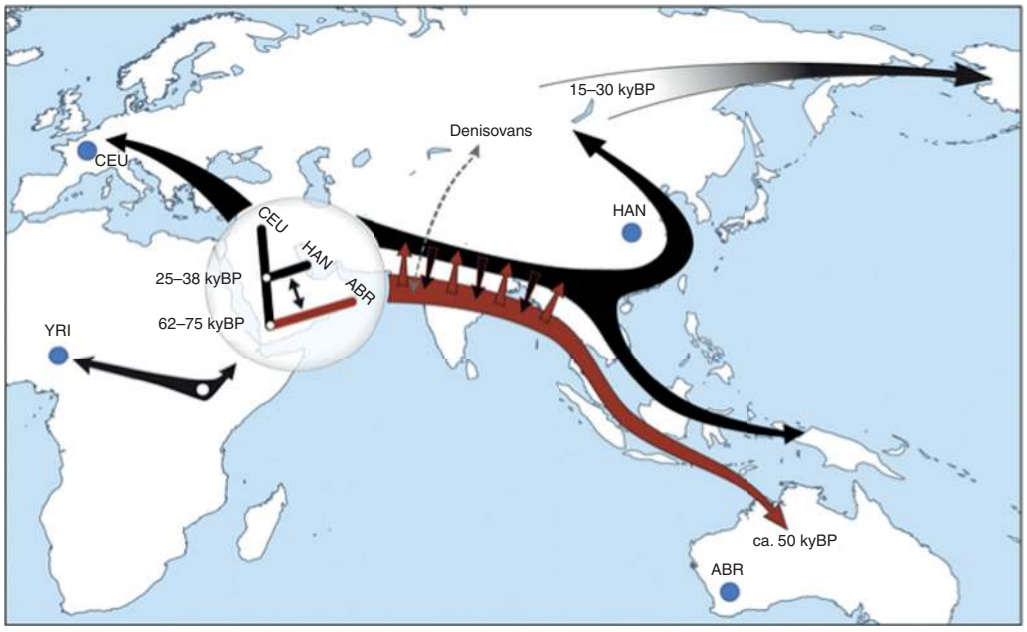


**Figure 5.** Reconstruction of early spread of modern humans outside Africa. YRI, Yoruba; ABR, Australian Aboriginal; CEU, European; HAN, Han Chinese; kyBP, thousand years before present. (From Rasmussen et al. 2011; with permission from the authors.)

ships and history (Paraskevis et al. 2013), the impact of significant population movements needs to be taken into account. For example, the slave trade in African peoples into South America over at least a 300-year period, and the more recent immigrations from Asia into Europe and North America in the last 200 years. Such population shifts introduce complexities that phylogenetic analysis cannot readily resolve. Therefore, characterization of the HBVs from remote and isolated indigenous peoples would provide a more accurate picture of human HBV origins.

We have assembled HBV isolates from indigenous populations including the Orang Asli from the Malaysian peninsula, indigenous people from the Torres Strait Islands (TSI), indigenous Australians from northern Australia, and Melanesians from Vanuatu. We also have retrieved from GenBank (Benson et al. 2005) HBV sequences from people of the Jarawas tribes who reside in the Andaman Islands of the Bay of Bengal (Murhekar et al. 2006), as well as sequences from Papua New Guinea. These populations are more likely to be direct descendants of the first wave of humans out of Africa. Remarkably, all of the HBV isolated from these populations was of genotype C (Fig. 6; Table 2). It should be noted that the HBV-C4 identified in indigenous Australians has, to date, been the exclusive strain isolated from that group (Davies et al. 2013). This HBV expresses the HBsAg serological subtype *ayw3*, which is an atypical serotype for genotype C HBVs. This unusual HBsAg of the C4 possibly arose as a consequence of a recombination event with the recently reported HBV genotype J from Borneo (Tatematsu et al. 2009; Littlejohn et al. 2014). Thus, it is reasonable to suggest that AMH were either infected with an ancestor strain of HBV genotype C when they left Africa or became infected as they migrated east along the coastal trail of the Indian ocean, out to the Andaman Islands, onto the Sunda shelf, into Sahul (Papua New Guinea and Australia), then Melanesia, and the Solomon Islands (Fig. 7), evolving into the very large number of HBV-C subgenotypes over time (Table 2) (Norder et al. 2004; Mulyanto et al. 2011, 2012).

However, there are at least two conundrums to be addressed. First, why is there no genotype C HBV reported in India or further west? Banerjee and colleagues (2006) showed a clear demarcation between genotype D in India/Bangladesh and the genotype C predominantly found in Myanmar/Thailand. A possible explanation is the climatic catastrophe following the Toba eruption in western Sumatra 73,000 $\pm$ 4000 years ago. The areas directly affected by ash fall from the Toba explosion included numerous sites in central India with evidence of a thickness of more than 6 meters (20 feet) (Acharyya and Basu 1993), and the associated "volcanic winter" has been estimated to have been at least 6 years in duration. The subsequent global cooling triggered by this 73 kya "super eruption" precipitated an environmental catastrophe that resulted in the near extinction of most contemporaneous human populations (Gibbons 1993; Hewitt 2000), but this devastation was confined to the West of Toba. Recent fossil evidence suggests that, East of Toba, mammals including gibbons and orangutans survived the immediate and postenvironmental changes of Toba and recovered to subsequently flourish (Louys 2007). This implies that any human populations on the Sunda shelf at that time would also have survived and quickly recovered, compared with their Indian subcontinent counterparts.

Second, why is there no HBV genotype C in Africa? Vartanian and colleagues (2002) discovered a novel strain of HBV (FG) isolated from a wild-caught chimpanzee (*Pan troglodytes schweinfurthii*) in East Africa. Subsequent molecular analysis of FG revealed that the strain is a recombinant between the genome of a ChHBV (backbone) and a 500-base-pair segment of an HBV-C2 genome (Magiorkinis et al. 2005). The C2 part of FG mapped closely with other C2 isolates recovered from the Asia-Pacific. It is significant that the chimpanzee was wild caught and from East Africa, bridging the Asian genotype C isolates with an African connection. Additional field studies are clearly needed to identify the broader African origins of genotype C–linked HBVs in and/or around Africa and their relationship to these indigenous strains along the beachcomber or coastal trail (Fig. 7).

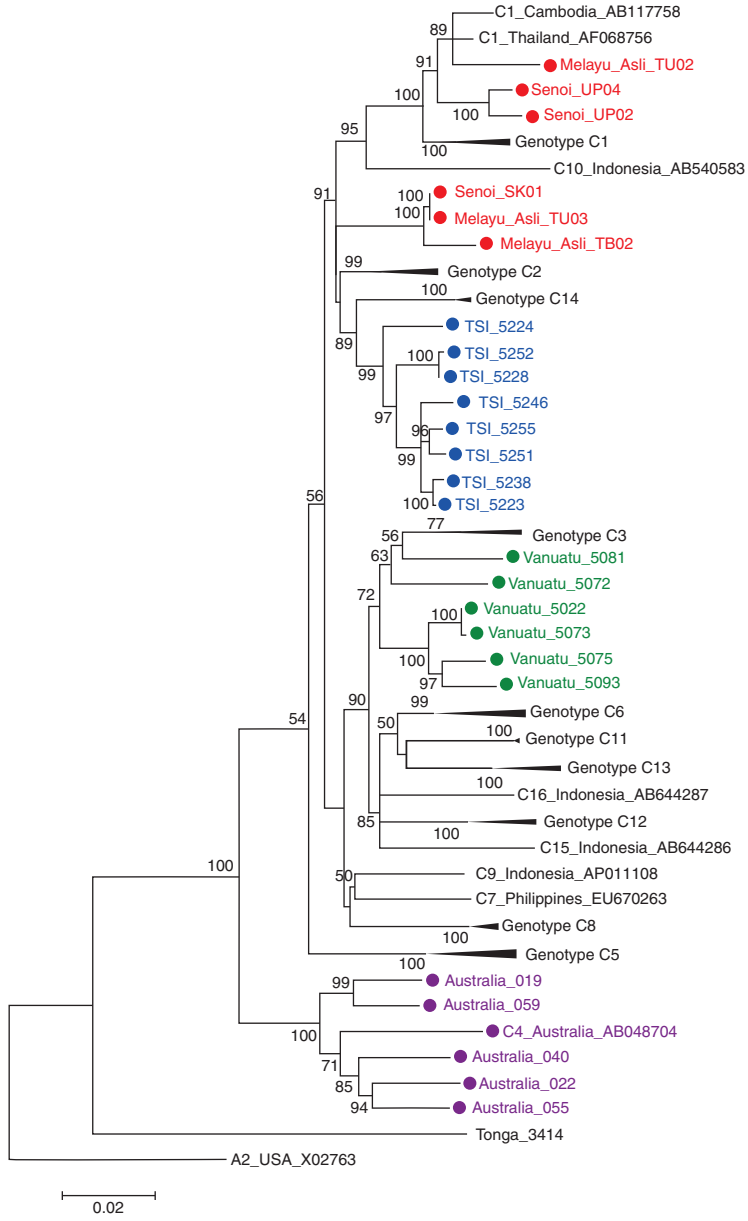


**Figure 6.** Phylogenetic tree of indigenous samples of HBV full genomes. The Orang Asli sequences form two groups (red): three clustered with C1 sequences, and three did not group with any other subgenotypes of C. The Torres Strait Islanders (blue) form their own clade, which shows a relationship with genotype C14. The Melanesians from Vanuatu clustered with C3 sequence, whereas the HBV from the indigenous Australians (purple) were genotype C4. The tree was generated using the maximum likelihood method. Numbers on branches indicate >50% bootstrap values (from 1000 resamplings).

**Table 2.** Summary of HBV genotypes and subgenotypes determined from the indigenous populations of East Africa along the beachcomber route

| Indigenous population | Geographical location | HBV subgenotype | HBV serotype | References |
|---|---|---|---|---|
| Jarawas Tribe | Andaman Island | C1 | *adr* | Murhekar et al. 2006 |
| Orang Asli | Malaysia | C1 | | N Aziz, M Littlejohn, L Yuen, et al., unpubl. |
| Torres Strait Islanders | Torres Strait Islands | C14-like | | V Ho, E Anderson, R Edwards, et al., unpubl. |
| Indigenous Australians | Northern Australia | C4 | *ayw3* | Sugauchi et al. 2001; Davies et al. 2013; Littlejohn et al. 2014 |
| Melanesians | Vanuatu | C3 | *adrq*$^-$ | Jazayeri et al. 2004 |

Based on these studies, we are proposing a model of HBV evolution, which suggests that modern-day HBVs are a result of multiple cross-species transmissions or zoonoses followed by subsequent recombination events on a genetic backbone of genotype C HBV infection in humans (Fig. 7).

## POSSIBLE ORIGINS OF HDV: HDV RNA AND HEPATITIS DELTA ANTIGEN (HDAg)

The HDV genome is a 1700-nucleotide negative-stranded circular RNA, which is ~70% self-complementary and forms a highly base-paired rod-like structure. It is comprised of

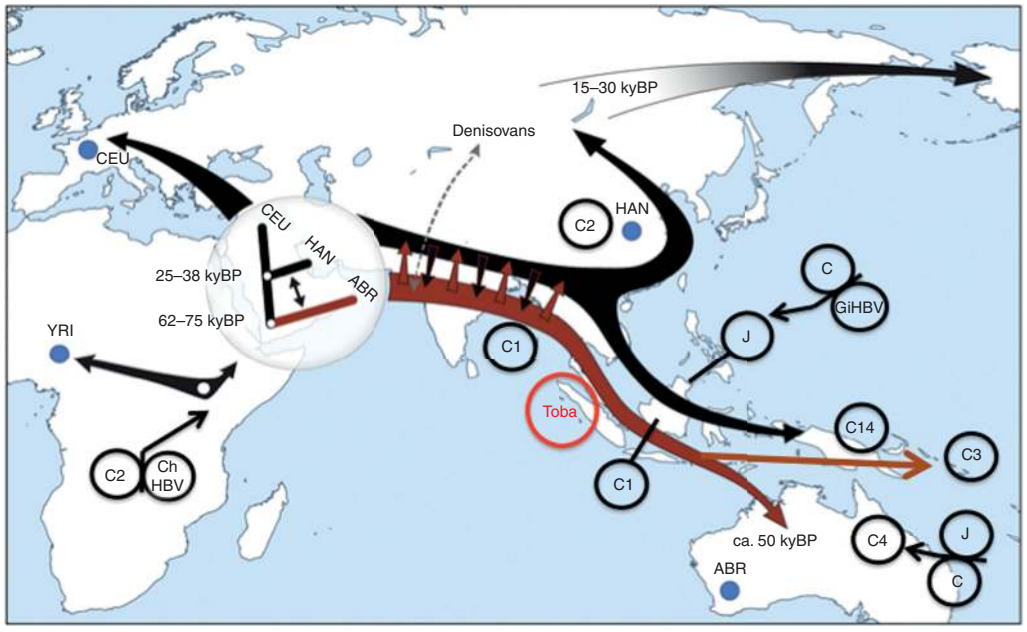


**Figure 7.** Reconstruction of early spread of modern humans outside Africa superimposed with the genotype C trail of HBV isolated from indigenous people along the coastal route. CEU, European; HAN, Han Chinese; ABR, Australian Aboriginal; YRI, Yoruba; GiHBV, gibbon HBV; ChHBV, chimpanzee HBV; kyBP, thousand years before present. (From Rasmussen et al. 2011; modified, with permission, from the authors.)

two separate domains: the viroid-like region encoding the self-cleaving ribozyme, and the protein-coding region encoding the HDAg. Based on nucleotide sequence analysis of available samples, HDV has been separated into eight genotypes, which generally have distinct geographical distributions. The most divergent HDVs have been isolated from African samples, suggesting a possible ancient African radiation (Radjef et al. 2004). A recent study of genotype 1 sequences from Turkey has suggested the presence of an amino acid polymorphism near the carboxyl terminus of large (L)-HDAg, which may be indicative of an African origin for sequences with a serine residue instead of an alanine at position 202 (Le Gal et al. 2012). Apart from these few studies, the evolution of HDV has not been well researched. Similarly, the origins of HDV are also unknown. The virus must have arisen in an HBV-infected cell due to the essential HBV helper function of HBsAg, which is required not only for HDV particle formation, but also for virion release and the next round of primary infections.

Brazas and Ganem (1996) proposed one of the first models for the origin of HDV, involving a two-step process in which a viroid-like RNA "captured" host-coding mRNAs. These investigators identified a cellular gene, termed delta-interacting protein A (DIPA), and they proposed that HDV RNA may have arisen when a free-living, self-replicating viroid like RNA "captured" a cellular mRNA encoding this DIPA protein or its common ancestor. It is worth noting that DIPA has nearly 60% amino acid similarity to HDAg.

Until recently, HDV was believed to be the only self-cleaving RNA associated with humans, albeit in the genome of a viral pathogen. At least two self-cleaving RNAs have now been found in mammals. The CLEC-2 ribozyme was identified in the mouse genome (Martick et al. 2008), whereas the cytoplasmic polyadenylation element-binding protein 3 (CPEB3) ribozyme is present in numerous mammalian genomes, including human (Salehi-Ashtiani et al. 2006). The ribozyme found in an intron in the CPEB3 gene is structurally and biochemically related to the HDV ribozyme, leading to the hypothesis that HDV actually arose directly from the human transcriptome (Salehi-Ashtiani et al. 2006). Because this ribozyme is found exclusively in mammals, it may have evolved as recently as 200 million years ago (Salehi-Ashtiani et al. 2006). The human CPEB3 ribozyme adopts a complex tertiary structure resembling the HDV ribozyme fold (Salehi-Ashtiani et al. 2006); however, there are no sequence similarities.

The third theory proposed for the origin of HDV comes from Taylor and colleagues who have based their model on the newly identified circular host mRNA molecules in cells (Taylor and Pelchat 2010; Taylor 2014). It would be from this pool of host RNA circles in hepatocytes (Memczak et al. 2013; Valdmanis and Kay 2013) that a rare RNA circle was selected caused by its ability to undergo RNA-directed replication using host cell RNA polymerase. The source, however, of the genetic information for the HDAg remains unknown, but HBV-spliced mRNAs or transcribed HBV DNA integrated sequences that have undergone extensive genetic drift could be possible candidates. Taylor and Pelchat (2010) have also pointed out that the ribozyme located on HDV antigenomic RNA is located downstream rather than upstream of the ORF for the HDAg. This downstream location is immediately preceded by polyadenylation sequence signals, such as AA UAAA, an arrangement typically used on host mRNAs (Hsieh and Taylor 1991). These investigators conclude that HDV may be no more than a selfish RNA. Whether or not it provides some function, negative or positive, in relation to its essential helper virus, HBV, the origins of HDV clearly warrants further investigation.

## SUMMARY AND FUTURE CONSIDERATIONS

A number of models for the origins of HBV have been proposed but none have become widely accepted because the conundrum of the viral mutation rate has not been adequately accounted for. Studies have consistently shown that genotype C is the oldest of the modern human HBVs and its HBsAg seems most closely related to primate HBsAg. We have described a

"coastal trail" of genotype C HBVs from Africa, to the Andaman Islands, onto the Sunda shelf, out into the Pacific, and also into Sahul (Fig. 7). Hepadnaviruses are clearly ancient viruses, having existed for at least 80 million years (Suh et al. 2013) and are well adapted to birds, bats, rodents, and higher primates. It would not be unreasonable to see the history of this family of viruses as a series of successful zoonoses from birds to rodents to bats to primates; elucidation of the evolutionary role that the NTCP receptor may play in this should be forthcoming in the very near future. These frequent cross-species transmissions have resulted in their widespread distribution in the animal kingdom and, as a result of a series of recombination events with other HBVs, both primate and human, resulted in the 10 genotypes of human HBVs that are recognized today. The human HBV genome can, thus, be thought to be made up of a number of modules of viral elements for regulation, structure, and function shaped by evolutionary pressures, resulting in the mosaic genome existing today (Fig. 3B) (Bowyer and Sim 2000; Simmonds and Midgley 2005). The public health implications of these cross-species transmissions and recombination events are potentially quite significant. This endemic infection in nonhuman primates has the potential for cross-species transmission, which will hamper control of HBV in geographical regions where humans and the higher primates share habitats. The detection of hepadnaviruses in bats adds another layer to the hepadnavirus story, highlighting the need for ongoing surveillance and monitoring for evidence of infection with these important oncogenic viruses.

In contrast, the origins and evolution of HDV remain to be adequately explained, and, at this point in time, a simple model of HDV evolution has yet to be proposed. Recent data has linked HDV and the plant viroids but do not resolve whether one is a precursor to the other or whether they represent convergent evolution. Recognition of circular forms of host RNAs suggests new insights into the origins of HDV and to the origins of life itself. More critical research is required to provide the data that will allow elucidation of the origins of this unique virus.

## REFERENCES

*Reference is also in this collection.


Acharyya S, Basu P. 1993. Toba ash on the Indian subcontinent and its implications for correlation of late pleistocene alluvium. *Quat Int* **40:** 10–19.

Ahn SH, Yuen L, Han KH, Littlejohn M, Chang HY, Damerow H, Ayres A, Heo J, Locarnini S, Revill PA. 2010. Molecular and clinical characteristics of hepatitis B virus in Korea. *J Med Virol* **82:** 1126–1134.

Banerjee A, Kurbanov F, Datta S, Chandra PK, Tanaka Y, Mizokami M, Chakravarty R. 2006. Phylogenetic relatedness and genetic diversity of hepatitis B virus isolates in Eastern India. *J Med Virol* **78:** 1164–1174.

Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL. 2005. GenBank. *Nucleic Acids Res* **33** (Database issue)**:** D34–D38.

Bollyky P, Rambaut A, Grassly NC, et al. 1997. Hepatitis B virus has a New World evolutionary origin. *Hepatology* **26:** 765.

Bowyer SM, Sim JG. 2000. Relationships within and between genotypes of hepatitis B virus at points across the genome: Footprints of recombination in certain isolates. *J Gen Virol* **81:** 379–392.

Brazas R, Ganem D. 1996. A cellular homolog of hepatitis delta antigen: Implications for viral replication and evolution. *Science* **274:** 90–94.

* Burns GS, Thompson AJ. 2014. Viral hepatitis B: Clinical and epidemiological characteristics. *Cold Spring Harb Perspect Med* **4:** a024935.

Calisher CH, Childs JE, Field HE, Holmes KV, Schountz T. 2006. Bats: Important reservoir hosts of emerging viruses. *Clin Microbiol Rev* **19:** 531–545.

Chan HL, Tsui SK, Tse CH, Ng EY, Au TC, Yuen L, Bartholomeusz A, Leung KS, Lee KH, Locarnini S, et al. 2005. Epidemiological and virological characteristics of 2 subgroups of hepatitis B virus genotype C. *J Infect Dis* **191:** 2022–2032.

Chen BF, Kao JH, Liu CJ, Chen DS, Chen PJ. 2004. Genotypic dominance and novel recombinations in HBV genotype B and C co-infected intravenous drug users. *J Med Virol* **73:** 13–22.

Chu DK, Peiris JS, Chen H, Guan Y, Poon LL. 2008. Genomic characterizations of bat coronaviruses (1A, 1B and HKU8) and evidence for co-infections in *Miniopterus* bats. *J Gen Virol* **89:** 1282–1287.

Cui J, Holmes EC. 2012. Endogenous hepadnaviruses in the genome of the budgerigar (*Melopsittacus undulatus*) and the evolution of avian hepadnaviruses. *J Virol* **86:** 7688–7691.

Cui C, Shi J, Hui L, Xi H, Zhuoma, Quni, Tesdan, Hu G. 2002. The dominant hepatitis B virus genotype identified in Tibet is a C/D hybrid. *J Gen Virol* **83:** 2773–2777.

Davies J, Littlejohn M, Locarnini SA, Whiting S, Hajkowicz K, Cowie BC, Bowden DS, Tong SY, Davis JS. 2013. Molecular epidemiology of hepatitis B in the Indigenous people of northern Australia. *J Gastroenterol Hepatol* **28:** 1234–1241.

Ding X, Gu H, Zhong ZH, Zilong X, Tran HT, Iwaki Y, Li TC, Sata T, Abe K. 2003. Molecular epidemiology of hepatitis

viruses and genotypic distribution of hepatitis B and C viruses in Harbin, China. *Jpn J Infect Dis* **56:** 19–22.

Dobson AP. 2005. Virology. What links bats to emerging infectious diseases? *Science* **310:** 628–629.

Drexler JF, Corman VM, Wegner T, Tateno AF, Zerbinati RM, Gloza-Rausch F, Seebens A, Müller MA, Drosten C. 2011. Amplification of emerging viruses in a bat colony. *Emerg Infect Dis* **17:** 449–456.

Drexler JF, Corman VM, Müller MA, Maganga GD, Vallo P, Binger T, Gloza-Rausch F, Cottontail VM, Rasche A, Yordanov S, et al. 2012. Bats host major mammalian paramyxoviruses. *Nat Commun* **3:** 796.

Drexler JF, Geipel A, Konig A, König A, Corman VM, van Riel D, Leijten LM, Bremer CM, Rasche A, Cottontail VM, et al. 2013. Bats carry pathogenic hepadnaviruses antigenically related to hepatitis B virus and capable of infecting human hepatocytes. *Proc Natl Acad Sci* **110:** 16151–16156.

Dupinay T, Gheit T, Roques P, Cova L, Chevallier-Queyron P, Tasahsu SI, Le Grand R, Simon F, Cordier G, Wakrim L, et al. 2013. Discovery of naturally occurring transmissible chronic hepatitis B virus infection among *Macaca fascicularis* from Mauritius Island. *Hepatology* **58:** 1610–1620.

Fischer SF, Schmidt K, Fiedler N, Glebe D, Schüttler C, Sun J, Gerlich WH, Repp R, Schaefer S. 2006. Genotype-dependent activation or repression of HBV enhancer II by transcription factor COUP-TF1. *World J Gastroenterol* **12:** 6054–6058.

Fujiwara K, Tanaka Y, Orito E, Ohno T, Kato T, Sugihara K, Hasegawa I, Sakurai M, Ito K, Ozasa A, et al. 2005. Distribution of HBV genotypes among HBV carriers in Benin: Phylogenetic analysis and virological characteristics of HBV genotype E. *World J Gastroenterol* **11:** 6410–6415.

Gibbons A. 1993. Pleistocene population explosions. *Science* **262:** 27–28.

Gibbons A. 2011. Anthropology. A new view of the birth of *Homo sapiens*. *Science* **331:** 392–394.

Gilbert C, Feschotte C. 2010. Genomic fossils calibrate the long-term evolution of hepadnaviruses. *PLoS Biol* **8:** e1000495.

Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, Patterson N, Li H, Zhai W, Fritz MH, et al. 2010. A draft sequence of the Neandertal genome. *Science* **328:** 710–722.

Grethe S, Heckel JO, Rietschel W, Hufert FT. 2000. Molecular epidemiology of hepatitis B virus variants in non-human primates. *J Virol* **74:** 5377–5381.

Hammer MF, Woerner AE, Mendez FL, Watkins JC, Wall JC. 2011. Genetic evidence for archaic admixture in Africa. *Proc Natl Acad Sci* **108:** 15123–15128.

Harrison A, Lemey P, Hurles M, Moyes C, Horn S, Pryor J, Malani J, Supuri M, Masta A, Teriboriki B, et al. 2011. Genomic analysis of hepatitis B virus reveals antigen state and genotype as sources of evolutionary rate variation. *Viruses* **23:** 83–101.

He B, Fan Q, Yang F, Hu T, Qiu W, Feng Y, Li Z, Li Y, Zhang F, Guo H, et al. 2013. Hepatitis virus in long-fingered bats, Myanmar. *Emerg Infect Dis* **19:** 638–640.

Hess G, Arnold W, Koesters W, Biswas R, Hütteroth TH, zum Büschenfelde KH. 1977. Simultaneous presence of HBsAg and anti-HBs in the serum of different subtypes (serological and immunofluorescent studies). *Z Immunitatsforsch Immunobiol* **153:** 143–151.

Hewitt G. 2000. The genetic legacy of the Quaternary ice ages. *Nature* **405:** 907–913.

Hsieh SY, Taylor J. 1991. Regulation of polyadenylation of hepatitis delta virus antigenomic RNA. *J Virol* **65:** 6438–6446.

Hu X, Margolis HS, Purcell RH, Ebert J, Robertson BH. 2000. Identification of hepatitis B virus indigenous to chimpanzees. *Proc Natl Acad Sci* **97:** 1661–1664.

Hu X, Javadian A, Gagneux P, Robertson BH. 2001. Paired chimpanzee hepatitis B virus (ChHBV) and mtDNA sequences suggest different ChHBV genetic variants are found in geographically distinct chimpanzee subspecies. *Virus Res* **79:** 103–108.

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

Huy TT, Ishikawa K, Ampofo W, Izumi T, Nakajima A, Ansah J, Tetteh JO, Nii-Trebi N, Aidoo S, Ofori-Adjei D, et al. 2006. Characteristics of hepatitis B virus in Ghana: Full length genome sequences indicate the endemicity of genotype E in West Africa. *J Med Virol* **78:** 178–184.

Huy TT, Ngoc TT, Abe K. 2008. New complex recombinant genotype of hepatitis B virus identified in Vietnam. *J Virol* **82:** 5657–5663.

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

Jazayeri SM, Alavian SM, Carman WF. 2010. Hepatitis B virus: Origin and evolution. *J Viral Hepat* **17:** 229–235.

Kahila Bar-Gal G, Kim MJ, Klein A, Shin DH, Oh CS, Kim JW, Kim TH, Kim SB, Grant PR, Pappo O, et al. 2012. Tracing hepatitis B virus to the 16th century in a Korean mummy. *Hepatology* **56:** 1671–1680.

Kao JH, Chen PJ, Lai MY, Chen DS. 2001. Acute exacerbations of chronic hepatitis B are rarely associated with superinfection of hepatitis B virus. *Hepatology* **34:** 817–823.

Kato H, Orito E, Sugauchi F, Ueda R, Koshizaka T, Yanaka S, Gish RG, Kurbanov F, Ruzibakiev R, Kramvis A, et al. 2003. Frequent coinfection with hepatitis B virus strains of distinct genotypes detected by hybridization with type-specific probes immobilized on a solid-phase support. *J Virol Methods* **110:** 29–35.

Kimbi GC, Kramvis A, Kew MC. 2004. Distinctive sequence characteristics of subgenotype A1 isolates of hepatitis B virus from South Africa. *J Gen Virol* **85:** 1211–1220.

Kramvis A, Paraskevis S. 2013. Subgenotype A1 of HBV-tracing human migrations in and out of Africa. *Antivir Ther* **18:** 513–521.

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

Kramvis A, Restorp K, Norder H, Botha JF, Magnius LO, Kew MC. 2005b. Full genome analysis of hepatitis B virus genotype E strains from South-Western Africa and Madagascar reveals low genetic variability. *J Med Virol* **77:** 47–52.

Lanford RE, Chavez D, Brasky KM, Burns RB III, Rico-Hesse R. 1998. Isolation of a hepadnavirus from the woolly monkey, a New World primate. *Proc Natl Acad Sci* **95:** 5757–5761.

Le Gal F, Badur S, Hawajri NA, Akyüz F, Kaymakoglu S, Brichler S, Zoulim F, Gordien E, Gault E, Dény P. 2012. Current hepatitis delta virus type 1 (HDV1) infections in central and eastern Turkey indicate a wide genetic diversity that is probably linked to different HDV1 origins. *Arch Virol* **157:** 647–659.

Littlejohn M, Davies J, Yuen L, Edwards R, Sozzi T, Jackson K, Cowie B, Tong S, Davis J, Locarnini S. 2014. Molecular virology of hepatitis B virus, sub-genotype C4 in Northern Australian indigenous populations. *J Med Virol* **86:** 695–706.

Louys J. 2007. Limited effect of the Quaternary's largest super-eruption (Toba) on land mammals from Southeast Asia. *Quat Sci Rev* **26:** 3108–3117.

MacDonald DM, Holmes EC, Lewis JC, Simmonds P. 2000. Detection of hepatitis B virus infection in wild-born chimpanzees (*Pan troglodytes verus*): Phylogenetic relationships with human and other primate genotypes. *J Virol* **74:** 4253–4257.

Magiorkinis EN, Magiorkinis GN, Paraskevis DN, Hatzakis AE. 2005. Re-analysis of a human hepatitis B virus (HBV) isolate from an East African wild born *Pan troglodytes schweinfurthii*: Evidence for interspecies recombination between HBV infecting chimpanzee and human. *Gene* **349:** 165–171.

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

Martick M, Horan LH, Noller HF, Scott WG. 2008. A discontinuous hammerhead ribozyme embedded in a mammalian messenger RNA. *Nature* **454:** 899–902.

Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A, Maier L, Mackowiak SD, Gregersen LH, Munschauer M, et al. 2013. Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* **495:** 333–338.

Miyakawa Y, Mizokami M. 2003. Classifying hepatitis B virus genotypes. *Intervirology* **46:** 329–338.

Motta-Castro AR, Martins RM, Yoshida CF, Teles SA, Paniago AM, Lima KM, Gomes SA. 2005. Hepatitis B virus infection in isolated Afro-Brazilian communities. *J Med Virol* **77:** 188–193.

Mulders MN, Venard V, Njayou M, Edorh AP, Bola Oyefolu AO, Kehinde MO, Muyembe Tamfum JJ, Nebie YK, Maiga I, Ammerlaan W, et al. 2004. Low genetic diversity despite hyperendemicity of hepatitis B virus genotype E throughout West Africa. *J Infect Dis* **190:** 400–408.

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

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

Murhekar MV, Chakravarty R, Murhekar KM, Banerjee A, Sehgal SC. 2006. Hepatitis B virus genotypes among the Jarawas: A primitive Negrito tribe of Andaman and Nicobar Islands, India. *Arch Virol* **151:** 1499–1510.

Newbold JE, Xin H, Tencza M, Sherman G, Dean J, Bowden S, Locarnini S. 1995. The covalently closed duplex form of the hepadnavirus genome exists in situ as a heterogeneous population of viral minichromosomes. *J Virol* **69:** 3350–3357.

Ni Y, Lempp F, Mehrle S, Nkongolo S, Kaufman C, Fälth M, Stindt J, Königer C, Nassal M, Kubitz R, et al. 2014. Hepatitis B virus exploits sodium taurocholate co-transporting polypeptide (NTCP) for entering hepatocytes in a species-specific manner. *Gastroenterology* **146:** 1070–1083.

Norder H, Courouce AM, Magnius LO. 1994. Complete genomes, phylogenetic relatedness, and structural proteins of six strains of the hepatitis B virus, four of which represent two new genotypes. *Virology* **198:** 489–503.

Norder H, Ebert JW, Fields HA, Mushahwar IK, Magnius LO. 1996. Complete sequencing of a gibbon hepatitis B virus genome reveals a unique genotype distantly related to the chimpanzee hepatitis B virus. *Virology* **218:** 214–223.

Norder H, Couroucé AM, Coursaget P, Echevarria JM, Lee SD, Mushahwar IK, Robertson BH, Locarnini S, Magnius LO. 2004. Genetic diversity of hepatitis B virus strains derived worldwide: Genotypes, subgenotypes, and HBsAg subtypes. *Intervirology* **47:** 289–309.

Okamoto H, Tsuda F, Sakugawa H, Sastrosoewignjo RI, Imai M, Miyakawa Y, Mayumi M. 1988. Typing hepatitis B virus by homology in nucleotide sequence: Comparison of surface antigen subtypes. *J Gen Virol* **69:** 2575–2583.

Olinger CM, Venard V, Njayou M, Oyefolu AO, Maïga I, Kemp AJ, Omilabu SA, le Faou A, Muller CP. 2006. Phylogenetic analysis of the precore/core gene of hepatitis B virus genotypes E and A in West Africa: New subtypes, mixed infections and recombinations. *J Gen Virol* **87:** 1163–1173.

Olinger CM, Jutavijittum P, Hubschen JM, Hübschen JM, Yousukh A, Samountry B, Thammavong T, Toriyama K, Muller CP. 2008. Possible new hepatitis B virus genotype, southeast Asia. *Emerg Infect Dis* **14:** 1777–1780.

Orito E, Mizokami M, Ina Y, Moriyama EN, Kameshima N, Yamamoto M, Gojobori T. 1989. Host-independent evolution and a genetic classification of the hepadnavirus family based on nucleotide sequences. *Proc Natl Acad Sci* **86:** 7059–7062.

Orito E, Ichida T, Sakugawa H, Sata M, Horiike N, Hino K, Okita K, Okanoue T, Iino S, Tanaka E, et al. 2001. Geographic distribution of hepatitis B virus (HBV) genotype in patients with chronic HBV infection in Japan. *Hepatology* **34:** 590–594.

Osiowy C, Giles E. 2003. Evaluation of the INNO-LiPA HBV genotyping assay for determination of hepatitis B virus genotype. *J Clin Microbiol* **41:** 5473–5477.

Paraskevis D, Magiorkinis G, Magiorkinis E, Ho SY, Belshaw R, Allain JP, Hatzakis A. 2013. Dating the origin and dispersal of hepatitis B virus infection in humans and primates. *Hepatology* **57:** 908–916.

Cold Spring Harbor Perspectives in Medicine
www.perspectivesinmedicine.org

Quintero A, Martinez D, Alarcón DeNoya B, Costagliola A, Urbina L, González N, Liprandi F, Castro De Guerra D, Pujol FH. 2002. Molecular epidemiology of hepatitis B virus in Afro-Venezuelan populations. *Arch Virol* **147:** 1829–1836.

Radjef N, Gordien E, Ivaniushina V, Gault E, Anaïs P, Drugan T, Trinchet JC, Roulot D, Tamby M, Milinkovitch MC, et al. 2004. Molecular phylogenetic analyses indicate a wide and ancient radiation of African hepatitis delta virus, suggesting a *Deltavirus* genus of at least seven major clades. *J Virol* **78:** 2537–2544.

Rasmussen M, Guo X, Wang Y, Lohmueller KE, Rasmussen S, Albrechtsen A, Skotte L, Lindgreen S, Metspalu M, Jombart T, et al. 2011. An Aboriginal Australian genome reveals separate human dispersals into Asia. *Science* **334:** 94–98.

Reich D, Green RE, Kircher M, Krause J, Patterson N, Durand EY, Viola B, Briggs AW, Stenzel U, Johnson PL, et al. 2010. Genetic history of an archaic hominin group from Denisova cave in Siberia. *Nature* **468:** 1053–1060.

Robertson BH, Margolis HS. 2002. Primate hepatitis B viruses—Genetic diversity, geography and evolution. *Rev Med Virol* **12:** 133–141.

Salehi-Ashtiani K, Lupták A, Litovchick A, Szostak JW. 2006. A genomewide search for ribozymes reveals an HDV-like sequence in the human CPEB3 gene. *Science* **313:** 1788–1792.

Sall AA, Starkman S, Reynes JM, Lay S, Nhim T, Hunt M, Marx N, Simmonds P. 2005. Frequent infection of *Hylobates pileatus* (pileated gibbon) with species-associated variants of hepatitis B virus in Cambodia. *J Gen Virol* **86:** 333–337.

Schlicht HJ, Salfeld J, Schaller H. 1987. The duck hepatitis B virus pre-C region encodes a signal sequence which is essential for synthesis and secretion of processed core proteins but not for virus formation. *J Virol* **61:** 3701–3709.

Seeger C, Baldwin B, Tennant BC. 1989. Expression of infectious woodchuck hepatitis virus in murine and avian fibroblasts. *J Virol* **63:** 4665–4669.

Sharp PM, Hahn BH. 2011. Origins of HIV and the AIDS pandemic. *Cold Spring Harb Perspect Med* **1:** a006841.

Simmonds P. 2001. Reconstructing the origins of human hepatitis viruses. *Philos Trans R Soc Lond B Biol Sci* **356:** 1013–1026.

Simmonds P, Midgley S. 2005. Recombination in the genesis and evolution of hepatitis B virus genotypes. *J Virol* **79:** 15467–15476.

Spradling P, Hu D, McMahon BJ. 2013. Epidemiology and prevention. In *Viral hepatitis*, 4th ed. (ed. Thomas H, et al.), pp. 81–95. Wiley, Hoboken, NJ.

Starkman SE, MacDonald DM, Lewis JC, Holmes EC, Simmonds P. 2003. Geographic and species association of hepatitis B virus genotypes in non-human primates. *Virology* **314:** 381–393.

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

Sugauchi F, Orito E, Ichida T, Kato H, Sakugawa H, Kakumu S, Ishida T, Chutaputti A, Lai CL, Ueda R, et al. 2002. Hepatitis B virus of genotype B with or without recombination with genotype C over the precore region plus the core gene. *J Virol* **76:** 5985–5992.

Sugauchi F, Kumada H, Acharya SA, Shrestha SM, Gamutan MT, Khan M, Gish RG, Tanaka Y, Kato T, Orito E, et al. 2004a. Epidemiological and sequence differences between two subtypes (Ae and Aa) of hepatitis B virus genotype A. *J Gen Virol* **85:** 811–820.

Sugauchi F, Kumada H, Sakugawa H, Komatsu M, Niitsuma H, Watanabe H, Akahane Y, Tokita H, Kato T, Tanaka Y, et al. 2004b. Two subtypes of genotype B (Ba and Bj) of hepatitis B virus in Japan. *Clin Infect Dis* **38:** 1222–1228.

Suh A, Brosius J, Schmitz J, Kriegs JO. 2013. The genome of a Mesozoic paleovirus reveals the evolution of hepatitis B viruses. *Nat Commun* **4:** 1791.

Tabor E, Gerety RJ, Smallwood LA, Barker LF. 1977. Coincident hepatitis B surface antigen and antibodies of different subtypes in human serum. *J Immunol* **118:** 369–370.

Takahashi K, Brotman B, Usuda S, Mishiro S, Prince AM. 2000. Full-genome sequence analyses of hepatitis B virus (HBV) strains recovered from chimpanzees infected in the wild: Implications for an origin of HBV. *Virology* **267:** 58–64.

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

Taylor JM. 2014. Host RNA circles and the origin of hepatitis delta virus. *World J Gastroenterol* **20:** 2971–2978.

Taylor J, Pelchat M. 2010. Origin of hepatitis delta virus. *Future Microbiol* **5:** 393–402.

Tedder RS, Bissett SL, Myers R, Ijaz S. 2013. The "Red Queen" dilemma—Running to stay in the same place: Reflections on the evolutionary vector of HBV in humans. *Antivir Ther* **18:** 489–496.

Valdmanis PN, Kay MA. 2013. The expanding repertoire of circular RNAs. *Mol Ther* **21:** 1112–1114.

Vartanian JP, Pineau P, Henry M, Hamilton WD, Muller MN, Wrangham RW, Wain-Hobson S. 2002. Identification of a hepatitis B virus genome in wild chimpanzees (*Pan troglodytes schweinfurthi*) from East Africa indicates a wide geographical dispersion among equatorial African primates. *J Virol* **76:** 11155–11158.

Wang Z, Liu Z, Zeng G, Wen S, Qi Y, Ma S, Naoumov NV, Hou J. 2005. A new intertype recombinant between genotypes C and D of hepatitis B virus identified in China. *J Gen Virol* **86:** 985–990.

Warren KS, Heeney JL, Swan RA, Heriyanto, Verschoor EJ. 1999. A new group of hepadnaviruses naturally infecting orangutans (*Pongo pygmaeus*). *J Virol* **73:** 7860–7865.

Yan H, Zhong G, Xu G, He W, Jing Z, Gao Z, Huang Y, Qi Y, Peng B, Wang H, et al. 2012. Sodium taurocholate cotransporting polypeptide is a functional receptor for human hepatitis B and D virus. *eLife* **1:** e00049.

Cold Spring Harbor Perspectives in Medicine
www.perspectivesinmedicine.org

Yang W, Summers J. 1995. Illegitimate replication of linear hepadnavirus DNA through nonhomologous recombination. *J Virol* **69:** 4029–4036.

Yang J, Xing K, Deng R, Wang J, Wang X. 2006. Identification of hepatitis B virus putative intergenotype recombinants by using fragment typing. *J Gen Virol* **87:** 2203–2215.

Zeng G, Wang Z, Wen S, Jiang J, Wang L, Cheng J, Tan D, Xiao F, Ma S, Li W, et al. 2005. Geographic distribution, virologic and clinical characteristics of hepatitis B virus genotypes in China. *J Viral Hepat* **12:** 609–617.

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



# Origins and Evolution of Hepatitis B Virus and Hepatitis D Virus

Margaret Littlejohn, Stephen Locarnini and Lilly Yuen

*Cold Spring Harb Perspect Med* 2016; doi: 10.1101/cshperspect.a021360

**Subject Collection** The Hepatitis B and Delta Viruses

**Hepatitis B Virus X and Regulation of Viral Gene Expression**
*Betty L. Slagle and Michael J. Bouchard*

**The Woodchuck, a Nonprimate Model for Immunopathogenesis and Therapeutic Immunomodulation in Chronic Hepatitis B Virus Infection**
*Michael Roggendorf, Anna D. Kosinska, Jia Liu, et al.*

**Mouse Models of Hepatitis B Virus Pathogenesis**
*Matteo Iannacone and Luca G. Guidotti*

**Therapy of Delta Hepatitis**
*Cihan Yurdaydin and Ramazan Idilman*

**Immune Response in Hepatitis B Virus Infection**
*Anthony Tan, Sarene Koh and Antonio Bertoletti*

**Hepatitis D Virus: Introduction and Epidemiology**
*Mario Rizzetto*

**Management of Chronic Hepatitis B in Patients from Special Populations**
*Ching-Lung Lai and Man-Fung Yuen*

**Hepatitis B Virus Genotypes and Variants**
*Chih-Lin Lin and Jia-Horng Kao*

**Origins and Evolution of Hepatitis B Virus and Hepatitis D Virus**
*Margaret Littlejohn, Stephen Locarnini and Lilly Yuen*

**Assembly and Release of Hepatitis B Virus**
*Lisa Selzer and Adam Zlotnick*

**Hepatitis D Virus Replication**
*John M. Taylor*

**Treatment of Liver Cancer**
*Chun-Yu Liu, Kuen-Feng Chen and Pei-Jer Chen*

**Hepatitis B Virus and Hepatitis D Virus Entry, Species Specificity, and Tissue Tropism**
*Koichi Watashi and Takaji Wakita*

**Hepadnavirus Genome Replication and Persistence**
*Jianming Hu and Christoph Seeger*

**The Chimpanzee Model for Hepatitis B Virus Infection**
*Stefan F. Wieland*

**Hepatitis B Virus Epidemiology**
*Jennifer H. MacLachlan and Benjamin C. Cowie*

For additional articles in this collection, see http://perspectivesinmedicine.cshlp.org/cgi/collection/