

Genotypes and Genetic Variability of Hepatitis B Virus

Anna Kramvis

Hepatitis Virus Diversity Research Programme (HVDRP), Department of Internal Medicine, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

Key Words

Phylogenetic analysis · Subgenotype · Serological subtype

Abstract

Sequence heterogeneity is a feature of hepatitis B virus (HBV), the prototype member of the family *Hepadnaviridae*. Based on an intergroup divergence of greater than 7.5% across the complete genome, HBV has been classified phylogenetically into 9 genotypes, A–I, with a putative 10th genotype 'J', isolated from a single individual. With between approximately 4 and 8% intergroup nucleotide divergence across the complete genome and good bootstrap support, genotypes A–D, F, H, and I are classified further into subgenotypes. There is a broad and highly statistically significant correlation between serological subtypes and genotypes, and in some cases, serological subtypes can be used to differentiate subgenotypes. The genotypes, and certain subgenotypes, have distinct geographical distributions and are important in both the clinical manifestation of infection and response to antiviral therapy. HBV genotypes/subgenotypes and genetic variability of HBV are useful in epidemiological and transmission studies, tracing human migrations, and in predicting the risk for the development of severe liver disease and response to antiviral therapy. Moreover, knowl-

edge of the genotype/subgenotype is important in implementing preventative strategies. Thus, it is crucial that new strains are correctly assigned to their respective genotype/subgenotype and consistent, unambiguous, and generally accepted nomenclature is utilized. © 2014 S. Karger AG, Basel

Hepatitis B virus (HBV), the prototype of the family *Hepadnaviridae*, has a partially double-stranded circular DNA genome of approximately 3,200 base pairs. This compact genome contains four partly or completely overlapping open reading frames: preC/C that encodes for e antigen (HBeAg) and core protein (HBcAg); P for polymerase (reverse transcriptase) (POL), S for surface proteins [three forms of HBsAg, small (S), middle (M) and large (L)], and X for a transcriptional transactivator protein. HBV replicates by reverse transcription of the pregenomic RNA, a 3.5-kb RNA intermediate, which is transcribed by the cellular RNA polymerase II from the covalently closed circular form of HBV DNA in the hepatocyte nucleus. Sequence heterogeneity is a feature of HBV because POL lacks proofreading ability. However, as a result of the constraints of the overlapping open reading frames and the presence of secondary RNA

structures, such as epsilon, coded by nonoverlapping regions, the mutation rate of the various regions of the HBV genome varies [1, 2]. The HBV genome has been estimated to evolve at an error rate of approximately 10^{-3} to 10^{-6} nucleotide substitutions/site/year [3–9]. To date, based on an intergroup divergence $>7.5\%$ across the complete genome, HBV has been classified phylogeneti-

cally into 9 genotypes, A–I [10–12], with a putative 10th genotype, 'J', isolated from a single individual [13]. With between approximately 4 and 8% intergroup nucleotide difference across the complete genome and good bootstrap support, genotypes A–D, F, H, and I are classified further into at least 35 subgenotypes (fig. 1; table 1) [11, 12, 14].

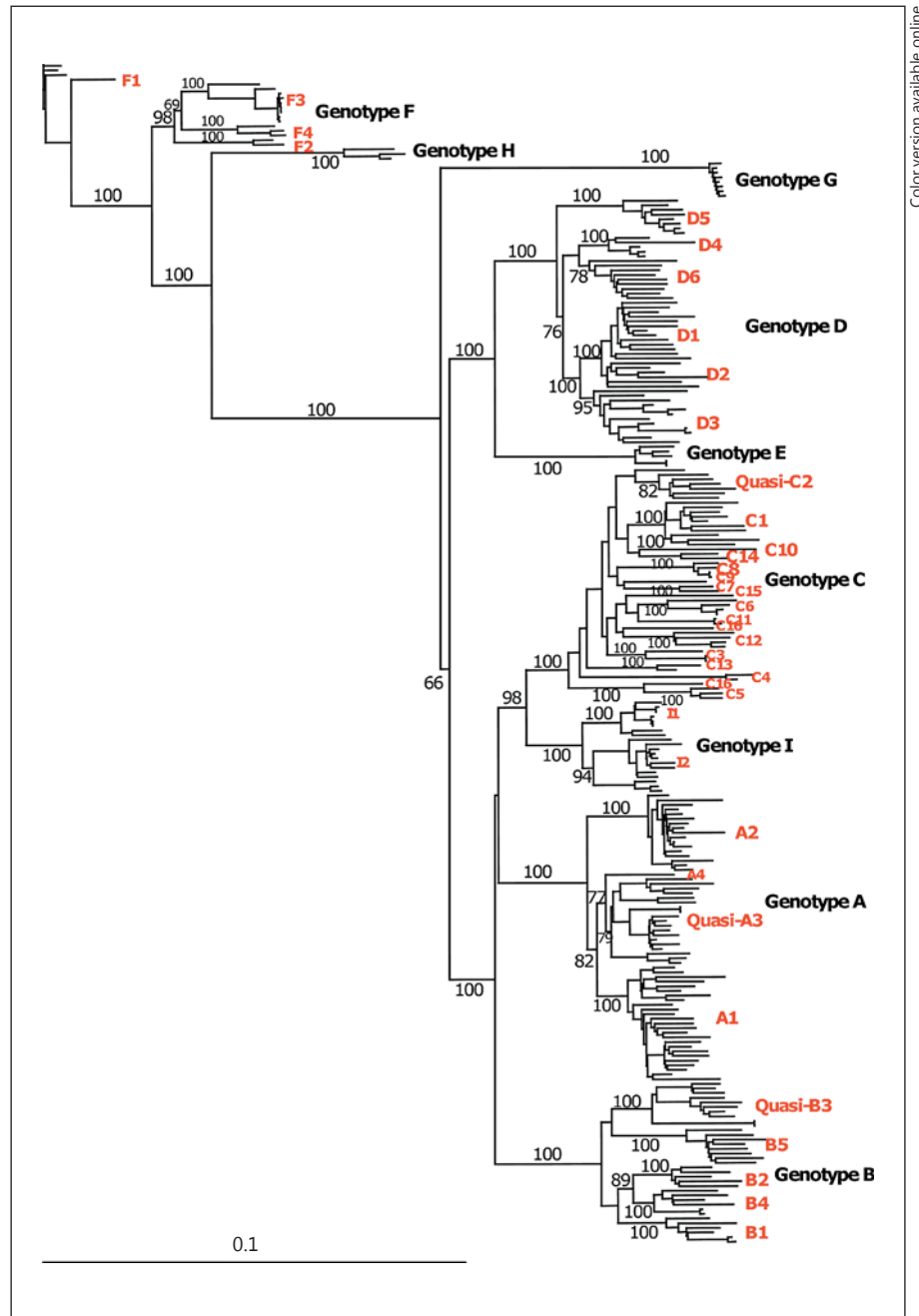


Fig. 1. A rooted phylogenetic tree of 250 full genome sequences of HBV representative of the 9 genotypes, established using neighbor joining. Bootstrap statistical analysis was performed using 1,000 data sets and the numbers on the nodes indicate the percentage of occurrences. Subgenotypes are shown in red (color refers to the online version only). GenBank accession numbers for the representative sequences included in the phylogenetic analyses are as follows:

(For second part of legend see next page.)

The genotypes differ in genome length (table 1), the size of open reading frames, and the proteins translated [11], as well as the development of various mutations [15]. Based on HBsAg heterogeneity [11], 9 serological subtypes, *ayw1*, *ayw2*, *ayw3*, *ayw4*, *ayr*, *adw2*, *adw4*, *adwq*, *adr*, and *adrq*-, have been identified. A broad, highly statistically significant correlation exists between

serological subtypes and genotypes: *adw* is associated with genotypes A, B, F, G, and H, *adr* with C, and *ayw* with D and E [14], but many exceptions exist. Moreover, mild immune selection may lead to mutations in the subtype determinants [W. Gerlich, pers. commun.]. In some cases, serological subtypes can be used to differentiate subgenotypes (table 1) [14]. Using the logistic regression

(Legend to Figure 1.)

Genotype A: subgenotype A1 (AB076678Malawi, AY233288South Africa, AB076679Malawi, FJ692592Haiti, AY233275-South Africa, AY233284South Africa, FM199974Rwanda, FM199979Rwanda, AY934765Somalia, AY934771Somalia, DQ020003UAE, AB116094Philippines, AB116093Philippines, AB116086India, AB453988Japan, AB453989Japan, AB116082-Bangladesh, AB116085Bangladesh, AB116088Nepal, AB116089-Nepal, M57663Philippines, AB116087India, AF090842Belgium, AF043560Argentina), subgenotype A2 (AY233280South Africa, AJ344115France, AY233286South Africa, Z35717Poland, Z72478Germany, AJ309370France, AF536524USA, AB014370-Japan, AY034878USA, X70185Germany, AY128092Canada, X02763USA, AB064314USA, AF537371USA, AF090838Belgium, AF090841Belgium, X51970Germany), quasi-subgenotype A3 (AY934763Gambia, AM180623Mali, AM180624Cameroon, AB194950Cameroon, AM184125Gambia, AM184126Gambia, AB194951Cameroon, AB194952Cameroon, FJ692554Nigeria, FJ692556Nigeria, FJ692607Haiti, FJ692600Haiti, FJ692604Haiti, FJ692602Haiti, FJ692606Haiti, FJ692601Haiti, FJ692603Haiti, FJ692605Haiti, AY934764Gambia, GQ161813Guinea), subgenotype A4 (GQ331048Belgium).

Genotype B: subgenotype B1 (AB010290Japan, AB010291-Japan, D00329Japan, AB073855 Japan, D23768Japan, D23679-Japan), subgenotype B2 (AB073831Thailand, AF282918China, AB073832Thailand, AB073836Taiwan, AB073839Taiwan), quasi-subgenotype B3 (D00331Indonesia, AB033555Sumatra, M54923Indonesia, AB219428Philippines, AB219427Philippines, AB241117Philippines, AB219426Philippines, AB219429-Philippines), subgenotype B4 (AY033072Netherlands*, AB115551Cambodia, AB117759Cambodia, AB073835Vietnam, AY033073Netherlands*, AB205122Vietnam), subgenotype B5 (DQ463797Canada, DQ463801Canada, DQ463799Canada, DQ463795Canada, DQ463802Canada, DQ463796Canada, DQ463798Canada, DQ463800Canada).

Genotype C: subgenotype C1 (AB112472Thailand, AB112471-Thailand, AB074756Thailand, AB117758Cambodia, AF068756-Thailand, AB112066Myanmar, AB112348Myanmar, AY167099-Taiwan, AB112063Vietnam, AB112065Vietnam), Quasi-subgenotype C2 (AY057947Tibet, AB014376Japan, AY066028China, AB014362Japan, AB014382Japan, S75184Japan, AB205123China, AY057947Tibet), subgenotype C3 (X75656Polynesia, AF458664-China, X75665New Caledonia), subgenotype C4 (AB048704-Australia, AB048705Australia), subgenotype C5 (AB241109-Philippines, AB241110Philippines, AB241111Philippines), subgenotype C6 (AB493842Indonesia, AB554014Indonesia,

AB493837Indonesia, AB554021Indonesia), subgenotype C7 (EU670263Philippines, GU721029South Korea), subgenotype C8 (AP011104Indonesia, AP011107Indonesia, AP011105Indonesia, AP011106Indonesia), subgenotype C9 (AP011108Indonesia), subgenotype C10 (AB540583Indonesia), subgenotype C11 (AB554020Indonesia, AB554019Indonesia), subgenotype C12 (AB554020Indonesia, AB554019Indonesia), subgenotype C12 (AB560662Indonesia, AB554025Indonesia, GQ358157Indonesia, AB554018Indonesia), subgenotype C13 (AB644280, AB644281), subgenotype C14 (GQ377555Indonesia, HM011493Indonesia), subgenotype C15 (AB644286Indonesia), subgenotype C16 (AB644287Indonesia).

Genotype D: subgenotype D1 (FJ899792China, AB104711-Egypt, JN642129Lebanon, GU456684Iran, AB583680Pakistan, FJ904415Tunisia, FJ386590China, GU456664Iran, GU456670-Iran, GU456673Iran, JF754607Turkey, JF754629Turkey), subgenotype D2 (GU456635Iran, EU594410Estonia, GQ477453-Poland, JF754597Turkey, JN642144Lebanon), subgenotype D3 (GQ922000Canada, EU594435Estonia, EU594434Estonia, AB493846Papua, AB493845Papua, AB493848Papua, AY233296-South Africa, U95551USA, AY902773USA, AY902776USA, FJ692507Haiti, X85254Italy), subgenotype D4 (AB033559Papua, AB048701Australia, FJ692536Haiti, FJ692532Haiti, FJ692533-Haiti), subgenotype D5 (AB033558Japan, GQ205382India, GQ205389India, GQ205384India, GQ205377India, GQ205385-India, GQ205378India, DQ355779India), subgenotype D6 (FJ904433Tunisia, FJ904441Tunisia, FJ904438 Tunisia, FJ904394Tunisia, FJ904395Tunisia, FJ904416Tunisia, FJ904398 Tunisia, FJ904407Tunisia, FJ904408Tunisia).

Genotype E: (X75664Senegal, AB091255Cote d'Ivoire, AB032431, AB091256Cote d'Ivoire, X75657Western Africa).

Genotype F: subgenotype F1 (AY090461El Salvador, AY090459-Costa Rica, AY090458Costa Rica, AF223963Argentina), subgenotype F2 (X69798Brazil, AY090455Nicaragua), subgenotype F4 (AF223965Argentina, AF223962Argentina, X75658France).

Genotype G: (AB056513USA, AB056514USA, AB056515USA, AB064310USA, AB064311USA, AB064312USA, AB064313USA, AF160501USA, AF405706Germany).

Genotype I: subgenotype I1 (AB231908Vietnam, AF241407-Vietnam, AF241408Vietnam, FJ023661Laos, FJ023660Laos, FJ023663Laos, AF241409Vietnam, GU357844China), subgenotype I2 (FJ023672Laos, EU833891Canada*, FJ023667Laos, FJ023664Laos, FJ023673Laos, FJ023665Laos, FJ023666Laos, FJ023669Laos, FJ023670Laos, EU835240India, EU835241India, EU835242India).

Table 1. Comparison of the virological and clinical characteristics of the genotypes of HBV^a

Geno- type	Length	Differentiating features	Subgenotypes	Geographic distribution	Serological subtype	Predominant transmission route
A	3,221	6-nucleotide insert at carboxyl end of core gene	A1 A2 quasi-subgenotype A3(A3, A4, A5) ^b A4(A6) ^b	Africa ^c Europe/North America Africa, Haiti	<i>adw2/ayw2</i> <i>adw2</i> <i>ayw1</i> <i>ayw1</i> <i>adw4</i> ^e	Horizontal: parenteral or sexual
B	3,215		B1 B2 quasi-subgenotype B3(B3, B5, B7–B9, B6 China) ^b B4 B5(B6) ^b	Japan China Indonesia/Philippines/China Vietnam/Cambodia/France Eskimos/Inuits	<i>adw2</i> <i>adw2</i> <i>adw2</i> <i>ayw1/adw2</i> <i>adw2</i> <i>adr</i> ^e	Perinatal
C	3,215		C1 quasi-subgenotype C2(C2, C14, undefined sequences) ^b C3 C4 C5 C6–C12 C13–C15 C16	Thailand/Myanmar/Vietnam Japan/China/Korea New Caledonia/Polynesia Australian Aborigines Philippines/Indonesia Indonesia/Philippines Indonesia Indonesia	<i>adr</i> <i>adr</i> <i>adr</i> <i>ayw2/ayw3</i> <i>adw2</i> <i>adr</i> <i>adr</i> <i>ayr</i> ^e	Perinatal
D	3,182	33-nucleotide deletion at the amino terminus of the preS1 region	D1 D2 D3 D4 D5 D6	Middle East, Central Asia Europe/Japan/Lebanon Worldwide Australian aborigines, Micronesians, Papua New Guineans, Arctic Denes India Tunisia/Nigeria	<i>ayw2</i> <i>ayw3</i> <i>ayw2/ayw3</i> <i>ayw2</i> <i>ayw3/ayw2</i> <i>ayw2</i> <i>ayw4</i> ^e / <i>adw3</i> ^e	Horizontal: parenteral with intravenous drug use being a risk factor
E	3,212	3-nucleotide deletion at the amino terminus of the preS1 region	–	Western/Central Africa	<i>ayw4</i>	Horizontal
F	3,215		F1 F2 F3 F4	Argentina/Costa Rica/El Salvador, Alaska Nicaragua/Venezuela/Brazil Venezuela/Colombia Argentina	<i>adw4</i> <i>adw4</i> <i>adw4</i> <i>adw4</i> <i>adw2</i> ^e / <i>ayw4</i> ^e	Horizontal
G	3,248	36-nucleotide insert, 3' of position 1905; 3-nucleotide deletion at the amino terminus of the preS1 region; stop codons at positions 2 and 28 of the precore region	–	USA/Mexico/ Germany/Italy/ UK/France	<i>adw2</i>	Horizontal: sexual transmission mainly in men- who-have-sex- with-men
H	3,215		–	Mexico/Japan/Nicaragua/USA	<i>adw4</i>	
I	3,215		I1 I2	Laos/Vietnam/ China Laos/India/Vietnam ^d	<i>adw2</i> <i>ayw2</i>	Perinatal
J	3,182	33-nucleotide deletion at the amino terminus of the preS1 region	–	Japan (possible origin Borneo)	<i>ayw3</i>	

^a Summarizes data compiled from [10, 11, 13, 14, 19, 25, 28] and references cited therein. ^b Earlier subgenotype designation. ^c And in regions outside Africa where there was historically forced migration as a result of the slave trade [18].

^d Vietnamese residing in Canada [43]. ^e Rare serological subtype for that genotype.

method, 1802–1803CG in the basic core promoter region was shown to be characteristic of genotypes A, D, and E, whereas 1802–1803TT is characteristic of genotypes B, C, and F; 1858C is positively associated with genotypes A, F, and H and 1858T with genotypes B, D, and E [14]. Subgenotype C2 has 1858T as opposed to C1 which has 1858C, and F1/F4 can be differentiated from F2/F3 by having 1858T instead of 1858C [14]. 1888A is positively associated with subgenotype A1 [14].

Genotype A and Its Subgenotypes

Genotype A is characterized by a 6-nucleotide insert at the carboxyl end of the core gene. Comprehensive analysis of genotype A [16, 17] has led to the classification of this genotype into subgenotypes A1, A2, A4, and quasi-subgenotype A3 because the latter group of sequences does not meet the criteria for a subgenotype classification [11, 14]. Subgenotypes A1, A4, and quasi-subgenotype A3 are found mainly in Africa, whereas A2 prevails in northern and central Europe and North America [18].

Genotype B and Its Subgenotypes

Using phylogenetics and sequence divergence of >4%, the subgenotypes of B have been reclassified into 6 subgenotypes: B1, B2, B4–B6, and quasi-subgenotype B3 (table 1) [19]. Subgenotype B1, found mainly in Japan [20], and B5 (previously B6) from a Canadian Inuit population [21] represent genotype B without recombination with genotype C in the precore/core region, as opposed to the remaining subgenotypes of B that have this recombination [20]. Subgenotype B1 was probably the ancestor of B5, possibly carried by indigenous peoples during migration from Siberia and Alaska to North America and Greenland [22, 23].

Genotype C and Its Subgenotypes

According to Paraskevis et al. [8] genotype C is the oldest HBV genotype. It has the highest number of subgenotypes, C1–C16 [24, 25], reflecting the long duration of its endemicity in humans. A large number of subgenotypes circulate in Indonesia [24]. Subgenotype C4 is exclusively found in indigenous people of northern Australia [26], who are descended from a founder group that emigrated from Africa at least 50,000 years ago [27].

Genotype D and Its Subgenotypes

In a recent systematic and comparative analysis of the subgenotypes of D, it was concluded that there are 6, not 8, subgenotypes. Subgenotypes D1–D6 can be differentiated by distinct clustering with high bootstrap support and signature amino acids. Subgenotypes D3 and ‘D6’ were reclassified as a single subgenotype D3, and ‘D8’ was shown to be a genotype D/E recombinant rather than a subgenotype [28]. Subgenotype D4, which is found in aboriginal populations in Papua New Guinea and Australia [12] and in a small percentage of the Canadian Inuit population [29], may be an early subgenotype dating from the time of early human intercontinental migrations. Moreover, a recombinant of subgenotype D4 has been identified in Sudan [30]. It is possible that D4 originated in Africa but has subsequently been replaced by other subgenotypes of D and the recombinant is a remnant of the original strain/s [30]. Subgenotype D4 carriers were significantly older than both D3 and B6 carriers [23]. Although it has been suggested that subgenotype D5 is the most ancient of the subgenotypes of D [31], this conclusion is neither supported by the relatively low intragroup divergence of this subgenotype [28], nor by the dated HBV phylogeny of the subgenotypes of D [8]. Additional sequences of subgenotype D5 from geographical regions besides the Paharia tribe of India [31] may resolve these discrepancies.

Genotype E

Genotype E has the unique serological subtype *ayw4* and can be differentiated from genotypes A–D, F, H, and I by a 3-nucleotide deletion in the preS1 region. This genotype is endemic in western, central Africa, with a low genetic diversity, intimating a recent emergence of 200 years or less [4]. As opposed to subgenotype A1, which was dispersed by the slave trade [18], genotype E is rarely found outside Africa, except in individuals of African descent, further supporting its recent emergence after the forced migrations of slavery. Using Bayesian inference, a median time of evolution from a most recent common ancestor (tMRCA) of 130 years has been calculated [4]. This differs from a tMRCA of 6,000 years estimated by others [8]. However, as suggested previously, it is possible that genotype E existed in indigenous African populations and has recently been reintroduced [32]. Genotype E has been isolated from

Pygmies [33] and the Khoi San [Kramvis, unpubl. data], and in Colombia [34] and India [35] in individuals with no history of travel to or from Africa. Without an accurate determination of the nucleotide substitution rate of HBV, the variance of the estimated age of genotype E will be difficult to resolve.

Genotype F and Its Subgenotypes

All genotype F isolates belong to serological subtype *adw4* and cluster into 4 subgenotypes, F1–F4. This genotype is found in the Amerindian populations of Central and South America, as well as Alaska.

Genotype G

Genotype G is characterized by a 36-nucleotide insert, 3' of position 1905, and two translational stop codons at positions 2 and 28 of the precore/core region, abrogating HBeAg expression [11]. This genotype can establish chronic infection only in the presence of other genotypes, most frequently genotype A, that can supply HBeAg *in trans*. A major risk factor is sexual transmission by men who have sex with men. Genotype G is least divergent from genotype E, with which it shares the 3-nucleotide deletion in the core region and a unique sequence in the preS [11, 36]. Although not yet detected in Africa, an African origin of genotype G has been postulated [36].

Genotype H

Genotype H prevails in Mexico in both the indigenous populations and mestizos (mixed descent), suggesting this genotype has a long history among the descendants of the Aztecs, before the arrival of Europeans [37]. It is most closely related to genotype F.

Genotype I and Its Subgenotypes

In 2008, sequence analysis of the complete genome of a single isolate (AB231908) from a Vietnamese male found it to be closely related to 3 previously described 'aberrant' Vietnamese strains [38] and a 9th genotype, I, was proposed [39]. This proposal was not accepted because the mean genetic divergence of these 4 strains

from genotype C was 7% and the recombination analysis was not robust [40]. Subsequently, sequences derived from Laos [41], the Idu Mishmi tribe in northeast India [42], a Canadian of Vietnamese descent [43], and China [10] have expanded the number of sequences. The nucleotide divergence of most of these sequences relative to genotype C was at least 7.5%, with good bootstrap support for the group, thus meeting the criteria for genotype assignment [14]. Two subgenotypes, I1 and I2, with serological subtypes *adw2* and *ayw2*, respectively, were described [41]. This separation into subgenotypes was questioned when additional strains from India clustering within subgenotype I2 were sequenced and the intersubgenotype divergence was calculated to be <4% [42]. By analyzing all 19 complete genotype I genomes, without indels, available in the GenBank, the intergroup divergence between subgenotype I1 and I2 was found to be $3.40 \pm 0.30\%$ (mean \pm SD), below the 4% cutoff. However, as for subgenotype D1 and D2 [28], an exception can be made because of the different serological subtypes. The highest intergroup divergence of 4.1% was between the Laotian strain FJ023663 (I1) and the Indian strain EU835242 (I2). The wide geographical distribution suggests this genotype has been endemic in a wide area of Asia for a long time [10]. Genotype I is a recombinant of genotypes A/C/G and an indeterminate genotype [10, 41–43], which clusters close to genotype C when the complete genome is analyzed, and with genotype A in the polymerase [10]. The genotype A and C regions are closely related to subgenotypes A3 and C3, respectively [10, 41–43]. Genotype I has been functionally characterized in both Huh7 cells and by acute hydrodynamic infection of a mouse. In both systems, genotype I secreted HBsAg at levels comparable to genotypes A, B, and C, and higher than D, but HBeAg at similar levels to genotype A but lower than B, C, and D [10].

'Genotype J'

This strain was isolated from a single Japanese man with hepatocellular carcinoma (HCC), who had lived in Borneo for a prolonged period of time [13]. The complete genome clusters with nonhuman HBV, including isolates from gibbons, orangutans, chimpanzees, and gorillas. When compared with 1,440 human and nonhuman HBV strains, its sequence diverged by 10.7–15.7% from other genotypes and did not show any evidence of recombination [13]. In a later analysis, using additional gibbon/

orangutan HBV sequences for comparison, Locarnini et al. [44] concluded that genotype 'J' is in fact a recombinant of genotype C and gibbon HBV in the S region. Thus, although the high intergroup divergence of genotype 'J' meets the criterion for classification into a separate genotype, it may represent a cross-species transmission [13], and identification and analysis of additional sequences will be required before the existence of this 10th genotype can be confirmed.

Recombinants

Variation and evolution of HBV has been influenced by recombination between genotypes [45]. By carrying out a recombination analysis of 3,400 nearly full-length HBV sequences, Shi et al. [46] identified 44 patterns of intergenotypic recombination. Only genotype H and putative genotype 'J' showed no recombination, whereas genotype I was composed entirely of recombinants and 93% of genotype B were recombinants [46]. The breakpoints occurred most frequently in the BCP/PC region, with breakpoints found in the small S and core regions [45, 46].

Geographical Distribution of Genotypes/ Subgenotypes

HBV genotypes, and in some cases subgenotypes, have distinct geographical distribution both globally and locally [11] (table 1). Genotype A circulates in Africa, Europe and the Americas, genotypes B and C in Asia, and genotypes F and H in Southern and Central America. Subgenotype A1 prevails in Africa and outside Africa in regions with a history of slave trade, whereas A2 is found mainly outside Africa [18]. Similarly, subgenotype B1 is generally confined to Japan and B2 is found in Southeast Asia. In the Canadian Arctic, genotype D prevails in the west whereas subgenotype B6 is found in the eastern regions [23]. These genotypes also have a distinct ethnic distribution, with genotype B found in the Inuits and genotype D in Denes [23]. The geographical distribution can be influenced by both ethnicity and migration between regions. In the USA, immigrants from Asia are infected with either genotype B or C, whereas in Europe, immigrants from Africa are generally infected with genotype E. Genotype I has been spread from Vietnam to Canada [43] and France [47] by immigration and adoption, respectively.

Effect of Genotypes/Subgenotypes on the Natural History of HBV Infection and Response to Antiviral Therapy

HBV genotypes may be responsible for differences in the natural history of chronic HBV infection [48], and therefore play a role in the clinical manifestation of infection and response to antiviral therapy [15, 49, 50]. Patients infected with genotypes A, B, D, and F show earlier and more frequent spontaneous HBeAg seroconversion compared to those infected with genotype C, independent of ethnicity [48, 51, 52]. Patients infected with genotype E have a higher frequency of HBeAg positivity and higher viral loads compared to patients infected with genotype D [30]. Individuals infected with subgenotype A1 lose HBeAg much earlier than those infected with subgenotype A2 [53]. Although HBeAg is not required for viral replication, it has been shown to act as an immune tolerogen [54] and to play a role in the transmission of HBV, viral persistence, and the establishment of chronic hepatitis B [55]. An essential requirement for acute HBV infection becoming chronic is the expression of HBeAg [56, 57]. In geographical regions with early HBeAg seroconversion, perinatal HBV transmission is rare because few women are HBeAg positive during their reproductive phase [58, 59]. Thus, perinatal transmission is common in areas where genotype C and I prevail (table 1). Transmission of HBV occurs in a higher proportion of children born to genotype C-infected mothers than genotype B-infected mothers [60].

HBV genotypes/subgenotypes A1, C, B2–B4, F1, and perhaps D show a higher risk of serious complications of HBV infection, including cirrhosis and development of HCC, compared to A2, B1, and B5 [50]. Genotype C infections have increased and earlier risk of liver inflammation, liver fibrosis, cirrhosis, and HCC than those infected with other genotypes [61]. The subgenotypes of B, with genotype C recombination, also cause more severe liver disease compared to subgenotypes of B without recombination. The risk of southern Africans infected with subgenotype A1 of developing HCC is 4.5 times that of patients infected with other genotypes, and they develop HCC 6.5 years earlier [62]. A similar association of subgenotype A1 with HCC and its development at a younger age is also found in southern Indians [63]. Subgenotype A2 is well adapted to sexual transmission and can establish chronic infection in adults [61]. In India, subgenotype D1 is significantly associated with chronic liver disease and D3 with occult HBV infection [64]. Alaskans infected with genotype C or F are more likely to revert to

HBeAg positivity and therefore are at more risk at developing severe liver disease than those infected with genotypes A, B, and D [52]. In Sudan, genotype E is found to prevail in asymptomatic blood donors [65], whereas genotype D predominates in liver disease patients [30]. Occult HBV infection, characterized by low viral loads, is common in patients infected with genotype H, and the development of HCC is rare [37].

Patients infected with genotype A or B respond better to interferon-based therapy compared to patients infected with either genotype C or D (reviewed in [15, 66]). No significant difference in response of the different genotypes/subgenotypes to nucleos(t)ide analogue therapy has been found [66].

The role of HBV genotypes in the clinical manifestation of disease following HBV infection may be related to the development of specific mutations in the BCP/PC region as well as deletions in the preS region of the genome. Genotype C has a higher propensity to develop mutations compared to genotype B [67]. For example, individuals infected with B5 (previously B6) did not develop HCC, which may be related to the fact that BCP/PC mutations and preS deletions do not develop in this subgenotype, whereas they were frequent in genotype F, which led to the development of HCC in infected individuals [68]. A discussion of the development of these mutations is beyond the scope of the present review and the reader is referred to the following papers and references cited therein [14, 15, 66, 67].

A number of studies have shown the genotypes/subgenotypes of HBV are important in both clinical manifestation of infection and response to antiviral therapy. HBV genotypes/subgenotypes and the genetic variability of HBV are useful in epidemiological studies, tracing human migrations, and in predicting the risk of development of severe liver disease and response to antiviral therapy. Moreover, knowledge of the genotype/subgenotype is important for implementing preventative strategies. Thus, it is crucial that new strains are correctly assigned to their respective genotype/subgenotype and that a consistent,

unambiguous, and generally accepted nomenclature is utilized. With the exponential growth in the number of sequences published in public databases, misclassifications have been described and corrected in some cases [16, 17, 19, 25, 69]. A set of minimum criteria should be followed when proposing a new genotype or subgenotype:

(1) A number of sequences of complete genomes of the correct length, for the respective genotype/subgenotype are essential [11, 70]; the clinical relevance and epidemiology of a genotype/subgenotype is impossible to determine from a single strain; single strains should be designated as a *putative* genotype/subgenotype until additional strains become available

(2) A nucleotide divergence of at least 7.5% to separate strains into genotypes [14]

(3) A nucleotide divergence of between approximately 4 and 7.5%, monophyletic clustering and good bootstrap support to separate subgenotypes [14]

(4) Distinct geographical separation (D1 vs. D2) and/or different serological subtypes (D1 vs. D2, and I1 vs. I2), monophyletic clustering and good bootstrap support, to separate subgenotypes, when the nucleotide divergence is <4%

(5) As many available sequences as possible, instead of a few representative ones, should be used for comparison, in order to make the analysis robust; the use of too few sequences can lead to misclassifications

(6) Sequences selected for comparison should not have indels or be recombinants of defined genotypes, as these can lead to an artificial increase in nucleotide divergence

(7) New strains should be checked for recombinations between genotypes/subgenotypes [46]

(8) Recent literature should be consulted in order to obtain up-to-date information on genotype/subgenotype assignment, misclassifications, or new assignments [16, 17, 19, 69]

This review is an update on three extensive and comprehensive analyses and reviews of 2005 [11, 15] and 2008 [14], and includes the newer proposals for subgenotype classification for genotype A [16, 17], B [19], C [25], and D [28].

References

- 1 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(Suppl 1):S83–S90.
- 2 Torres C, Fernandez MD, Flichman DM, Campos RH, Mbayed VA: Influence of overlapping genes on the evolution of human hepatitis B virus. *Virology* 2013;441:40–48.
- 3 Tedder RS, Bissett SL, Myers R, Ijaz S: The 'Red Queen' dilemma – running to stay in the same place: reflections on the evolutionary vector of HBV in humans. *Antivir Ther* 2013; 18:489–496.
- 4 Andernach IE, Hunewald OE, Müller CP: Bayesian inference of the evolution of HBV/E. *PLoS One* 2013;8:e81690.

- 5 Fares MA, Holmes EC: A revised evolutionary history of hepatitis B virus (HBV). *J Mol Evol* 2002;54:807–814.
- 6 Orito E, Mizokami M, Ina Y, Moriyama EN, Kameshima N, Yamamoto M, Gojobori T: Host-independent evolution and a genetic classification of the hepadnavirus family based on nucleotide sequences. *Proc Natl Acad Sci USA* 1989;86:7059–7062.
- 7 Günther S, Sommer G, Von Breunig F, Iwanska A, Kalinina T, Sterneck M, Will H: Amplification of full-length hepatitis B virus genomes from samples from patients with low levels of viremia: frequency and functional consequences of PCR-introduced mutations. *J Clin Microbiol* 1998;36:531–538.
- 8 Paraskevis D, Magiorkinis G, Magiorkinis E, Ho SY, Belshaw R, Allain JP, Hatzakis A: Dating the origin and dispersal of hepatitis B virus infection in humans and primates. *Hepatology* 2013;57:908–916.
- 9 Bouckaert R, Alvarado-Mora MV, Pinho JR: Evolutionary rates and HBV: issues of rate estimation with Bayesian molecular methods. *Antivir Ther* 2013;18:497–503.
- 10 Yu H, Yuan Q, Ge SX, Wang HY, Zhang YL, Chen QR, Zhang J, Chen PJ, Xia NS: Molecular and phylogenetic analyses suggest an additional hepatitis B virus genotype 'T'. *PLoS One* 2010;5:e9297.
- 11 Kramvis A, Kew M, Francois G: Hepatitis B virus genotypes. *Vaccine* 2005;23:2409–2423.
- 12 Norder H, Courouce AM, Coursaget P, Echevarria JM, Lee SD, Mushahwar IK, Robertson BH, Locarnini S, Magnius LO: Genetic diversity of hepatitis B virus strains derived worldwide: genotypes, subgenotypes, and HBsAg subtypes. *Intervirology* 2004;47:289–309.
- 13 Tatematsu K, Tanaka Y, Kurbanov F, Sugauchi F, Mano S, Maeshiro T, Nakayoshi T, Wakuta M, Miyakawa Y, Mizokami M: 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.
- 14 Kramvis A, Arakawa K, Yu MC, Nogueira R, Stram DO, Kew MC: Relationship of serological subtype, basic core promoter and precore mutations to genotypes/subgenotypes of hepatitis B virus. *J Med Virol* 2008;80:27–46.
- 15 Kramvis A, Kew MC: Relationship of genotypes of hepatitis B virus to mutations, disease progression and response to antiviral therapy. *J Viral Hepat* 2005;12:456–464.
- 16 Pourkarim MR, Amini-Bavil-Olyae S, Lemey P, Maes P, Van Ranst M: Are hepatitis B virus 'subgenotypes' defined accurately? *J Clin Virol* 2010;47:356–360.
- 17 Pourkarim MR, Amini-Bavil-Olyae S, Lemey P, Maes P, Van Ranst M: HBV subgenotype misclassification expands quasi-subgenotype A3. *Clin Microbiol Infect* 2011;17:947–949.
- 18 Kramvis A, Paraskevis D: Subgenotype A1 of HBV – tracing human migrations in and out of Africa. *Antivir Ther* 2013;18:513–521.
- 19 Shi W, Zhu C, Zheng W, Carr MJ, Higgins DG, Zhang Z: Subgenotype reclassification of genotype B hepatitis B virus. *BMC Gastroenterol* 2012;12:116.
- 20 Sugauchi F, Orito E, Ichida T, Kato H, Sakugawa H, Kakumu S, Ishida T, Chutaputti A, Lai CL, Ueda R, Miyakawa Y, Mizokami M: Hepatitis B virus of genotype B with or without recombination with genotype C over the pre-core region plus the core gene. *J Virol* 2002;76:5985–5992.
- 21 Osiowy C, Giles E, Tanaka Y, Mizokami M, Minuk GY: Molecular evolution of hepatitis B virus over 25 years. *J Virol* 2006;80:10307–10314.
- 22 Sakamoto T, Tanaka Y, Simonetti J, Osiowy C, Borresen ML, Koch A, Kurbanov F, Sugiyama M, Minuk GY, McMahon BJ, Joh T, Mizokami M: Classification of hepatitis B virus genotype B into 2 major types based on characterization of a novel subgenotype in Arctic indigenous populations. *J Infect Dis* 2007;196:1487–1492.
- 23 Osiowy C, Larke B, Giles E: Distinct geographical and demographic distribution of hepatitis B virus genotypes in the Canadian Arctic as revealed through an extensive molecular epidemiological survey. *J Viral Hepat* 2011;18:e11–e19.
- 24 Mulyanto, Pancawardani P, Depamede SN, Wahyono A, Jirintai S, Nagashima S, Takahashi M, Nishizawa T, Okamoto H: 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* 2012;163:129–140.
- 25 Shi W, Zhu C, Zheng W, Ling C, Carr MJ, Higgins DG, Zhang Z: Subgenotyping of genotype C hepatitis B virus: correcting misclassifications and identifying a novel subgenotype. *PLoS One* 2012;7:e47271.
- 26 Davies J, Littlejohn M, Locarnini SA, Whiting S, Hajkowicz K, Cowie BC, Bowden DS, Tong SY, Davis JS: Molecular epidemiology of hepatitis B in the indigenous people of northern Australia. *J Gastroenterol Hepatol* 2013;28:1234–1241.
- 27 Hudjashov G, Kivisild T, Underhill PA, Endicott P, Sanchez JJ, Lin AA, Shen P, Oefner P, Renfrew C, Villems R, Forster P: Revealing the prehistoric settlement of Australia by Y chromosome and mtDNA analysis. *Proc Natl Acad Sci USA* 2007;104:8726–8730.
- 28 Yousif M, Kramvis A: Genotype D of hepatitis B virus and its subgenotypes: an update. *Hepatology Res* 2013;43:355–364.
- 29 Osiowy C, Larke B, Giles E: Distinct geographical and demographic distribution of hepatitis B virus genotypes in the Canadian Arctic as revealed through an extensive molecular epidemiological survey. *J Viral Hepat* 2011;18:e11–e19.
- 30 Yousif M, Mudawi H, Bakhiet S, Glebe D, Kramvis A: Molecular characterization of hepatitis B virus in liver disease patients and asymptomatic carriers of the virus in Sudan. *BMC Infect Dis* 2013;13:328.
- 31 Ghosh S, Banerjee P, RoyChoudhury A, Sarkar S, Ghosh A, Santra A, Banerjee S, Das K, Dwivedi B, Kar SK, Rao VG, Bhat JT, Singh N, Chowdhury A, Datta S: Unique hepatitis B virus subgenotype in a primitive tribal community in eastern India. *J Clin Microbiol* 2010;48:4063–4071.
- 32 Kramvis A, Kew MC: Epidemiology of hepatitis B virus in Africa, its genotypes and clinical associations of genotypes. *Hepatology Res* 2007;37:S9–S19.
- 33 Kurbanov F, Tanaka Y, Fujiwara K, Sugauchi F, Mbanya D, Zekeng L, Ndemi N, Ngansop C, Kaptue L, Miura T, Ido E, Hayami M, Ichimura H, Mizokami M: A new subtype (subgenotype) Ac (A3) of hepatitis B virus and recombination between genotypes A and E in Cameroon. *J Gen Virol* 2005;86:2047–2056.
- 34 Alvarado Mora MV, Romano CM, Gomes-Gouveia MS, Gutierrez MF, Carrilho FJ, Pinho JR: Molecular epidemiology and genetic diversity of hepatitis B virus genotype E in an isolated Afro-Colombian community. *J Gen Virol* 2010;91:501–508.
- 35 Singh J, Dickens C, Pahal V, Kumar R, Chaudhary R, Kramvis A, Kew MC: First report of genotype e of hepatitis B virus in an Indian population. *Intervirology* 2009;52:235–238.
- 36 Lindh M: HBV genotype G – an odd genotype of unknown origin. *J Clin Virol* 2005;34:315–316.
- 37 Panduro A, Maldonado-Gonzalez M, Fierro NA, Roman S: Distribution of HBV genotypes F and H in Mexico and Central America. *Antivir Ther* 2013;18:475–484.
- 38 Hannoun C, Norder H, Lindh M: An aberrant genotype revealed in recombinant hepatitis B virus strains from Vietnam. *J Gen Virol* 2000;81:2267–2272.
- 39 Huy TT, Trinh TN, Abe K: New complex recombinant genotype of hepatitis B virus identified in Vietnam. *J Virol* 2008;82:5657–5663.
- 40 Kurbanov F, Tanaka Y, Kramvis A, Simmonds P, Mizokami M: When should 'T' consider a new hepatitis B virus genotype? *J Virol* 2008;82:8241–8242.
- 41 Olinger CM, Jutavittum P, Hubschen JM, Yousukh A, Samouny B, Thammavong T, Toriyama K, Muller CP: Possible new hepatitis B virus genotype, southeast Asia. *Emerg Infect Dis* 2008;14:1777–1780.
- 42 Arankalle VA, Gandhe SS, Borkakoty BJ, Walimbe AM, Biswas D, Mahanta J: A novel HBV recombinant (genotype I) similar to Vietnam/Laos in a primitive tribe in eastern India. *J Viral Hepat* 2010:501–510.
- 43 Osiowy C, Kaita K, Solar K, Mendoza K: Molecular characterization of hepatitis B virus and a 9-year clinical profile in a patient infected with genotype I. *J Med Virol* 2010;82:942–948.
- 44 Locarnini S, Littlejohn M, Aziz MN, Yuen L: Possible origins and evolution of the hepatitis B virus (HBV). *Semin Cancer Biol* 2013;23:561–575.

- 45 Simmonds P, Midgley S: Recombination in the genesis and evolution of hepatitis B virus genotypes. *J Virol* 2005;79:15467–15476.
- 46 Shi W, Carr MJ, Dunford L, Zhu C, Hall WW, Higgins DG: Identification of novel inter-genotypic recombinants of human hepatitis B viruses by large-scale phylogenetic analysis. *Virology* 2012;427:51–59.
- 47 Colson P, Roquelaure B, Tamalet C: Detection of a newly identified hepatitis B virus genotype in southeastern France. *J Clin Virol* 2009;45:165–167.
- 48 Liaw YF, Brunetto MR, Hadziyannis S: The natural history of chronic HBV infection and geographical differences. *Antivir Ther* 2010; 15(suppl 3):25–33.
- 49 Kao JH: Hepatitis B viral genotypes: clinical relevance and molecular characteristics. *J Gastroenterol Hepatol* 2002;17:643–650.
- 50 McMahon BJ: The influence of hepatitis B virus genotype and subgenotype on the natural history of chronic hepatitis B. *Hepatol Int* 2009;3:334–342.
- 51 Chu CJ, Hussain M, Lok AS: Hepatitis B virus genotype B is associated with earlier HBeAg seroconversion compared with hepatitis B virus genotype C. *Gastroenterology* 2002;122: 1756–1762.
- 52 Livingston SE, Simonetti JP, Bulkow LR, Homan CE, Snowball MM, Cagle HH, Negus SE, McMahon BJ: Clearance of hepatitis B e antigen in patients with chronic hepatitis B and genotypes A, B, C, D, and F. *Gastroenterology* 2007;133:1452–1457.
- 53 Tanaka Y, Hasegawa I, Kato T, Orito E, Hirashima N, Acharya SK, Gish RG, Kramvis A, Kew MC, Yoshihara N, Shrestha SM, Khan M, Miyakawa Y, Mizokami M: A case-control study for differences among hepatitis B virus infections of genotypes A (subtypes Aa and Ae) and D. *Hepatology* 2004;40:747–755.
- 54 Chen M, Sallberg M, Hughes J, Jones J, Guidotti LG, Chisari FV, Billaud JN, Milich DR: Immune tolerance split between hepatitis B virus precore and core proteins. *J Virol* 2005;79:3016–3027.
- 55 Milich D, Liang TJ: Exploring the biological basis of hepatitis B e antigen in hepatitis B virus infection. *Hepatology* 2003;38:1075–1086.
- 56 Cote PJ, Korba BE, Miller RH, Jacob JR, Baldwin BH, Hornbuckle WE, Purcell RH, Tennant BC, Gerin JL: Effects of age and viral determinants on chronicity as an outcome of experimental woodchuck hepatitis virus infection. *Hepatology* 2000;31:190–200.
- 57 Hadziyannis SJ, Vassilopoulos D: Hepatitis B e antigen-negative chronic hepatitis B. *Hepatology* 2001;34:617–624.
- 58 Hadziyannis SJ: Natural history of chronic hepatitis B in Euro-Mediterranean and African countries. *J Hepatol* 2011;55:183–191.
- 59 Okada K, Kamiyama I, Inomata M, Imai M, Miyakawa Y: e antigen and anti-e in the serum of asymptomatic carrier mothers as indicators of positive and negative transmission of hepatitis B virus to their infants. *N Engl J Med* 1976;294:746–749.
- 60 Wen WH, Chen HL, Ni YH, Hsu HY, Kao JH, Hu FC, Chang MH: Secular trend of the viral genotype distribution in children with chronic hepatitis B virus infection after universal infant immunization. *Hepatology* 2011;53: 429–436.
- 61 Araujo NM, Waizbort R, Kay A: Hepatitis B virus infection from an evolutionary point of view: how viral, host, and environmental factors shape genotypes and subgenotypes. *Infect Genet Evol* 2011;11:1199–1207.
- 62 Kew MC, Kramvis A, Yu MC, Arakawa K, Hodgkinson J: Increased hepatocarcinogenic potential of hepatitis B virus genotype A in Bantu-speaking sub-Saharan Africans. *J Med Virol* 2005;75:513–521.
- 63 Gopalakrishnan D, Keyter M, Shenoy KT, Leena KB, Thayumanavan L, Thomas V, Vinayakumar K, Panackel C, Korah AT, Nair R: Hepatitis B virus subgenotype A1 predominates in liver disease patients from Kerala, India. *World J Gastroenterol* 2013;19:9294–9306.
- 64 Chandra PK, Biswas A, Datta S, Banerjee A, Panigrahi R, Chakrabarti S, De BK, Chakravarty R: Subgenotypes of hepatitis B virus genotype D (D1, D2, D3 and D5) in India: differential pattern of mutations, liver injury and occult HBV infection. *J Viral Hepat* 2009;16: 749–756.
- 65 Mahgoub S, Candotti D, El Ekiaby M, Allain JP: Hepatitis B virus (HBV) infection and recombination between HBV genotypes D and E in asymptomatic blood donors from Khartoum, Sudan. *J Clin Microbiol* 2011;49:298–306.
- 66 Lin CL, Kao JH: Clinical implications of hepatitis B virus variants. *J Formos Med Assoc* 2010;109:321–325.
- 67 Liu S, Zhang H, Gu C, Yin J, He Y, Xie J, Cao G: Associations between hepatitis B virus mutations and the risk of hepatocellular carcinoma: a meta-analysis. *J Natl Cancer Inst* 2009; 101:1066–1082.
- 68 Kowalec K, Minuk GY, Borresen ML, Koch A, McMahon BJ, Simons B, Osiowy C: Genetic diversity of hepatitis B virus genotypes B6, D and F among circumpolar indigenous individuals. *J Viral Hepat* 2013;20:122–130.
- 69 Shi W, Zhang Z, Ling C, Zheng W, Zhu C, Carr MJ, Higgins DG: Hepatitis B virus subgenotyping: history, effects of recombination, misclassifications, and corrections. *Infect Genet Evol* 2013;16:355–361.
- 70 Kurbanov F, Tanaka Y, Mizokami M: Geographical and genetic diversity of the human hepatitis B virus. *Hepatol Res* 2010;40:14–30.